MIT

FLIGHT TRANSPORTATION LABORATORY
REPORT R 90-5

AIRLINE RESERVATIONS FORECASTING: PROBABILISTIC AND STATISTICAL MODELS OF THE BOOKING PROCESS

Anthony O. Lee

September 1990

DEPARTMENT
OF
AERONAUTICS
&
ASTRONAUTICS

FLIGHT TRANSPORTATION
LABORATORY
Cambridge, Mass. 02139

AIRLINE RESERVATIONS FORECASTING:

PROBABILISTIC AND STATISTICAL MODELS OF THE BOOKING PROCESS

by

ANTHONY OWEN LEE

Sc.B. Applied Mathematics, Brown University (1984)

S.M. Operations Research, Massachusetts Institute of Technology (1987)

SUBMITTED TO THE DEPARTMENT OF CIVIL ENGINEERING
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY
in Transportation Systems

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
September, 1990

Signature of Author..
Department of Civil Engineering
August, 1990

Certified by..
Moshe Ben-Akiva
Professor, Civil Engineering
Thesis Supervisor

Accepted by..
Ole S. Madsen
Departmental Committee on Graduate Studies

AIRLINE RESERVATIONS FORECASTING:

PROBABILISTIC AND STATISTICAL MODELS OF THE BOOKING PROCESS

by

ANTHONY OWEN LEE

Submitted to the Department of Civil Engineering
in August, 1990 in partial fulfillment of the requirements
for the Degree of Doctor of Philosophy in Transportation Systems

ABSTRACT

In this thesis, we develop the necessary statistical framework to produce accurate forecasts of total bookings in a particular fare class on a specific flight number departing on a given date at various points before departure. After an introduction to the basic terminology of the airline booking process, a rigorous probabilistic model is developed. The booking process is modeled as a stochastic process with requests, reservations, and cancellations interspersed in the time before a flight departs. The key result of the probabilistic analysis is a censored Poisson model of the airline booking process.

A comprehensive statistical framework views the booking process from a data analysis perspective. We describe models based on advance bookings (the traditional booking curve) and historical bookings (a traditional time series model). An important development is the combined model which features a potentially more accurate combination of the advance bookings and historical bookings models. Additionally, we extend the statistical framework to include booking limits, which constrain the observed number of reservations in each fare class. The result is a truncated-censored regression model with truncation from below at zero and censoring from above at the booking limit.

We test the forecasting ability of the censored Poisson model and a combined statistical model with censored Normal errors using actual airline data provided by a major U.S. airline. When compared to industry standard models, the models developed in this thesis produce significant improvements in forecast accuracy. In the appendix, a Monte Carlo simulation is performed to determine the value of accurate forecasting for the airlines. The results demonstrate that each 10% improvement in forecast accuracy can bring about a 0.5% to 3.0% increase in expected revenues.

Thesis Supervisor: Dr. Moshe Ben-Akiva

Title: Professor of Civil Engineering

ACKNOWLEDGEMENTS

"Sow for yourselves righteousness; Reap the fruit of unfailing love; and break up your unplowed ground, for it is time to seek the Lord until He comes and showers His righteousness upon you." -- Hosea 10:12.

I would like to thank my thesis supervisor, Professor Moshe Ben-Akiva, for all the help and encouragement that he has given me over the past 3 years. It has truly been challenging working under such a well-respected person in the field of demand modeling. Moshe encouraged me to step out on faith and formulate the probabilistic model presented in this thesis.

My project supervisor, Professor Peter Belobaba, has been a great help in learning the "ins" and "outs" of revenue management and seat inventory control. His intuition and knowledge about airline management were quite helpful in the preparation of this thesis. He managed to keep this research down to earth and relevant to the airline industry.

I would like to thank Prof. Bob Simpson for contributing his vast knowledge of air transportation to this research effort. Also, I owe a great deal to Prof. Haris Koutsopoulos for reading this thesis and for encouraging me during this entire research effort.

I would also like to thank all the various students in CTS, FTL, and TSD for their support during the preparation of this thesis. Specifically, Biz Williamson and Grace Zabat took time to listen to me when nobody else would. Also, all of my office mates (Ted, Scott, Shangyao, Dharma, and Isam) were generous in their time and patience. Staff member Lisa Magnano was extremely helpful during the final weeks of preparation of this thesis. Thanks, Lisa!

My parents and grandparents have been extremely generous and supportive during my entire decade of higher education. I sincerely thank them for all of their encouragement, help, and support. Without them, I would not have been here.

Finally, I dedicate this thesis in memory of two of my favorite people: Professor Stella Dafermos and my grandfather, Mr. Frank Little. Both of them played special roles in my life.

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Motivation for Thesis

The need for accurate reservations forecasting in the U.S. airline industry is a direct result of the U.S. Airline Deregulation Act of 1978. In 1978, the United States government gave up authority over domestic fares and routes, allowing U.S. airlines to enter and leave domestic markets freely and to charge whatever fare the market would bear. Two immediate results of deregulation were increased competition on many routes and pricing freedom. In the late 1970's and early 1980's, the entrance of new, low-cost airlines into traditionally profitable markets threatened the established, high-cost carriers. Often, the entrance of new airlines into a market brought about substantially lower fares. In order to remain profitable, airline managers of the established carriers recognized the need to match the lower fares of the newer carriers on a limited basis, while still retaining many seats for sale at higher fares.

Pricing freedom, the second key result of deregulation, allowed airlines to charge whatever fare the market would bear. Airlines were able to change their pricing structure to take into account the fundamental differences between leisure travelers and business travelers. Leisure travelers tend to price sensitive and book well before the day of departure. By contrast, business travelers are generally time sensitive and make reservations closer to the day of departure. The airlines began to offer low fares with many restrictions to price sensitive leisure travelers while maintaining higher fares with fewer, if any, restrictions for the time sensitive business traveler. In order to protect space for the higher paying, late booking business travelers, airline managers needed to limit the space available to the early booking, low fare paying leisure travelers.

Thus, on two dimensions, the U.S. airlines recognized the need to limit seats available to lower fare paying passengers and protect seats for high fare paying passengers. In essence, the airlines desire to closely control their inventory of seats, making intelligent decisions of how many seats to sell in each fare class. The process of *seat inventory control* (also called "yield management") attempts to find the optimal allocation of spaces to each fare class by maximizing revenue. With the deregulation of Canadian domestic airlines, the planned liberalization of European markets in 1992, and the effect of the U.S. "free market" philosophy in international aviation treaties, the interest in seat inventory control is no longer limited primarily to U.S. airlines but is spreading to many airlines throughout the world.

The goal of the seat inventory control process is to maximize revenue by optimally allocating the seats on the aircraft among the various fare classes. One of the most critical inputs of an effective seat inventory control system is historical reservations data. Then, for future flight departures, the seat inventory control process:

1. Produces accurate forecasts of the total number of bookings in each fare class.
2. Optimizes the allocation of seats among fare classes, given the forecasts of total bookings.

The output of a seat inventory control system is the optimum allocation of seats for each fare class on a future departure. Figure 1.1 illustrates a seat inventory control system.

Much research has been done on the second step, the optimal allocation of seats among fare classes. However, relatively little attention has been paid to airline reservations forecasting. In Appendix A, we attempt to quantify the potential benefits of accurate forecasting methods. A

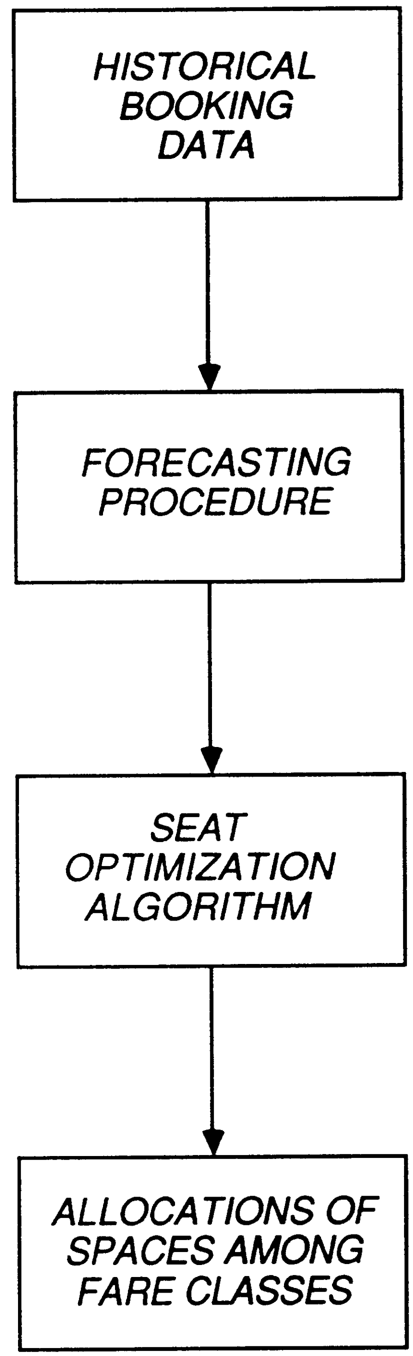

Figure 1.1 The Seat Inventory Control Process

simulation of the change in expected revenues brought about by more accurate forecasts is performed. The results demonstrate that increased forecast accuracy does indeed bring about increased revenues, particularly on high demand flights. Furthermore, we show that each 10% improvement in forecast accuracy on high demand flights can potentially result in a $10 to $60 million increase in total annual revenue for a major U.S. airline. Thus, the benefits of pursuing accurate forecasting methods are quite clear. This thesis will address many of the fundamental issues related to accurate airline reservations forecasting.

1.2 Goal of Thesis

The goal of this thesis is to develop the necessary statistical framework to produce accurate forecasts of total bookings in a particular fare class on a specific flight number departing on a given date at various times before departure. In order to attain this goal, we have three primary objectives in this thesis. First, we develop a rigorous probabilistic model, which describes the booking process as a stochastic process with requests, reservations, and cancellations interspersed in the time before departure of a flight. The probabilistic model emphasizes the dynamic nature of the airline booking process. The key results of the probabilistic analysis are a relatively straightforward model based on the Poisson distribution and a starting point for developing simulations of the airline booking process.

The second objective is to create a comprehensive statistical framework from the data analysis perspective. We introduce models based on advance bookings (the traditional booking curve) and historical bookings (time series of bookings). An important development is the concept of a combined model, which combines the advance bookings and historical bookings models to produce a potentially more accurate forecasting model. In addition, we introduce the concept of booking limits, which constrain the observed number of reservations in each fare class. When the booking limits enter into the analysis, the combined model is extended to a truncated-censored combined model with truncation from below at zero and censoring from above at capacity. We investigate a truncated-censored combined model based on the Normal distribution.

After the development of a probabilistic model and a rigorous statistical framework, the third objective of this thesis is to demonstrate the effectiveness of the probabilistic and statistical models on actual airline data. When compared to industry standard models, we show that the models developed in this thesis can produce significant improvements in forecast accuracy. The results validate the application of these probabilistic and statistical methods to airline reservations forecasting.

1.3 Organization of Thesis

The rest of this thesis is organized in the following manner. Chapter 2, entitled *Basic Definitions and Economic Analysis of the Booking Process*, introduces the terminology of the airline booking process in the context of a microeconomic framework. The booking process is a series of interactions between a potential air traveler and an airline. Detailed descriptions of the three phases of the booking process, which include the reservation phase, the cancellation phase, and the boarding phase, are given. The second part of the chapter focuses specifically on a microeconomic framework for the airline booking process. In microeconomic terms, the booking process is an interaction between a utility-maximizing consumer and a profit-maximizing airline. We formulate the profit maximization problem which the airline faces and the utility maximization problem that the consumer confronts. Then, we examine the forum in which airline reservations are made, specifically discussing the roles of the travel agent and the computer reservations system. Finally, the booking data which results from the interaction between the consumer and the airline is described.

Chapter 3, *Previous Approaches: A Literature Review,* surveys the relevant literature devoted to airline forecasting problems. We point out that, while most of the literature discusses macro-level models or passenger choice models, the area of primary interest in this thesis is micro-level forecasting. First, we examine the area of macro-level forecasting, which is the most heavily studied area. Macro-level forecasting includes aggregate forecasts such as the annual number of U.S. to Europe passengers or the number of passengers who board at a certain airport. Second, we briefly detail the small amount of work which has been done on passenger choice modeling. Finally, we discuss the literature on micro-level forecasts. This literature examines forecasting at the fare class and flight level. However, we find that the scope of the previous work in micro-level forecasting is quite limited.

Chapter 4, entitled *A Probabilistic Model of the Booking Process,* analyzes the booking process from a probabilistic standpoint. First, we discuss the assumptions of stationary and distinct airline booking data. Next, the airline booking process is described as a stochastic process. This stochastic process includes no-shows, go-shows, and waitlists, usually regarded as special cases, in a straightforward and natural manner. After simplifying assumptions are made, the booking process is modeled as an immigration and death process. Then, this chapter discusses the applicability of stochastic models to forecasting and the practical issues of estimation and forecasting for stochastic models. Chapter 4 concludes with a description of the "ideal" stochastic model, which incorporates the more subtle aspects of the booking process.

The topic of Chapter 5 is the development of a comprehensive statistical framework for analysis of the airline booking process. Chapter 5 examines the booking process from the statistical data analysis viewpoint. First, we investigate the available airline data on-hand at any time during the booking process of a particular flight. Then, we make the fundamental distinction between estimation, which fits a model to historical data, and forecasting, which uses the estimated model to predict future bookings. Three types of statistical models naturally arise from the data: advance bookings, historical bookings, and combined models. We show that the combined models provide an intuitive view of the booking process. The second part of Chapter 5 addresses the issues of booking limits and demand distributions. First, the effect of booking limits on airline booking data is considered. The presence of booking limits leads to a truncated-censored demand distribution and a set of truncated-censored regression equations. The chapter concludes with a discussion of the strengths and weaknesses of several possible demand distributions, such as the Normal, Log-normal, Poisson, and Gamma.

Chapter 6, *Practical Issues in Estimation and Forecasting*, addresses the practical issues involved in estimation and forecasting of the airline booking process. First, the types of booking data generally available to airlines are outlined. Also, we discuss the types of data not generally available which would be of value in analyzing the airline booking process. This chapter continues with the important issues in the estimation and forecasting of airline bookings. For example, we investigate issues such as the amount of historical data to use in estimation, the frequency of re-estimating the models, the detection and removal of outliers from the booking data, the effect of seasonal variation, and model selection. Key issues in forecasting include how to measure forecast performance and how to avoid "bad" forecasts. Two case studies illustrate the importance of the above issues in estimation and forecasting.

Chapter 7, entitled *Model Estimation and Forecasting*, involves the formulation and testing of probabilistic and statistical models on actual airline booking data provided by a major U.S. airline. We formulate the likelihood functions of a statistical model based on the Normal distribution and a probabilistic model based on the Poisson distribution. Two case studies apply the models to actual booking data and report significant improvements in forecasting accuracy over simple linear regression and moving average models.

Chapter 8 contains conclusions and topics of further research in airline reservations forecasting. The chapter begins with a summary of the major contributions and results of this thesis. Then, extensions of the probabilistic and statistical models are discussed. We briefly address the issue of origin-destination forecasting. Finally, this chapter examines the practical issues surrounding the implementation of an airline reservations forecasting system.

Chapter 2 Basic Definitions and Economic Analysis of the Booking Process

2.1 Introduction

This chapter describes the airline booking process in the context of a microeconomic framework. First, the booking process is described in detail. The next section introduces the microeconomic analysis of the booking process, focusing on the supply of air transportation by the airline firm and the generation of demand by the consumer. The computer reservations system (CRS) is the forum in which the transactions between the consumer and the airline take place and are recorded. Finally, this chapter examines the interaction between the consumer and the firm in light of their respective goals and demonstrates the role of forecasting from the airline's perspective.

2.2 A Description of the Airline Booking Process

This section follows an individual reservation through the booking process and defines the relevant terms. The booking process consists of three phases: the reservation phase, the cancellation phase, and the boarding phase. In the *reservation phase*, a potential air traveler makes a request for air travel and the airline attempts to satisfy the request. In the *cancellation phase*, air travel plans may be cancelled during the time before the flight departs. The final part of the booking process is the *boarding phase*, which occurs at the airport on the day of the flight. Figure 2.1 illustrates the entire booking process.

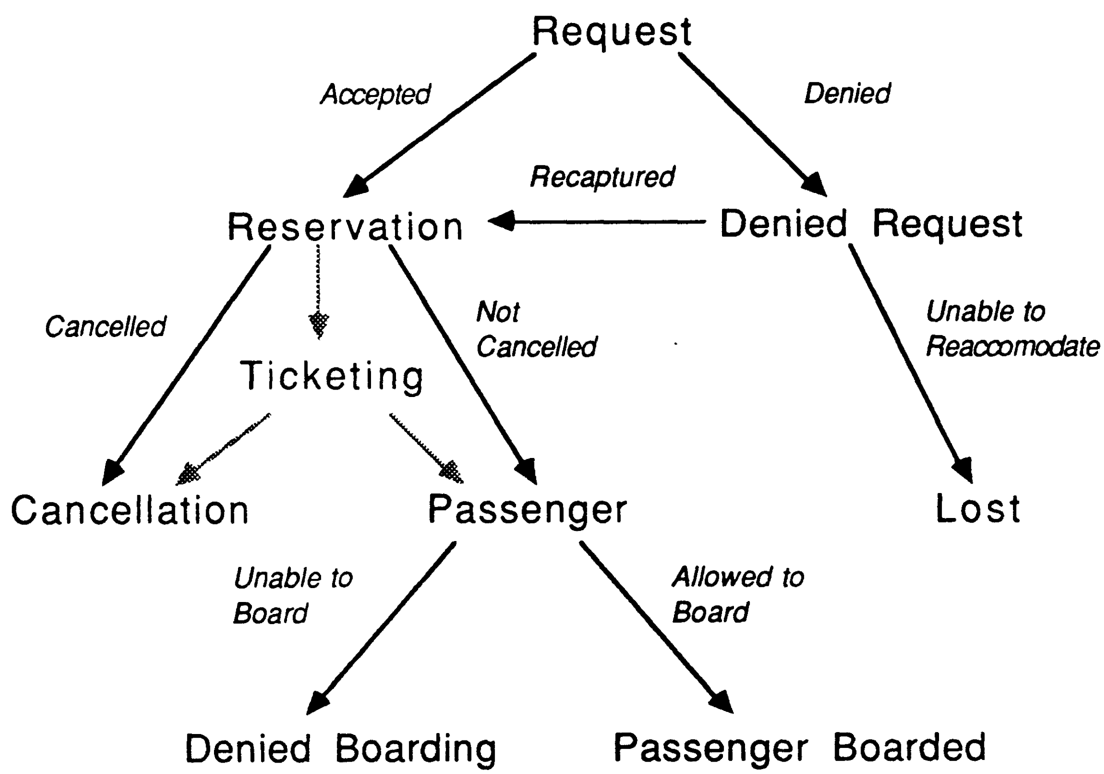

Figure 2.1 The Booking Process

2.2.1 The Reservation Phase

In the reservation phase, a request for air travel services enters the computer reservations system. A *request* is defined as a call to a travel agent or airline reservations agent for air travel services between an origin city and a destination city on a certain date. A request usually includes return service from the destination city back to the origin city at a later time. For the purposes of this section, we assume that a request is for a given level of service on a specific itinerary, which consists of one or more flights (aircraft flying between a city pair at a specific time) at a particular fare (the price of travel between the origin and destination cities) on specified dates. Requests are usually made by a potential traveler in the days and weeks before the desired departure date. However, a request may be made at the airport on the day of departure. A *go-show* is a traveler who shows up at the airport on the day of departure and makes a last minute request for air travel on a particular flight.

If space is available on the desired flight at the desired fare, a reservation is made. A *reservation* is an accepted request for a *space* on a specific flight at a particular fare on a given date. It is important to draw the distinction between a space and a seat on an aircraft. An airline may sell more spaces than physical seats, because some travelers holding reservations may not show up on the day of departure. A flight is said to be *overbooked* when the total number of reservations exceeds the actual number of seats on the aircraft.

On the other hand, if space is not available between the origin and destination cities on the requested date at the desired fare, the request is denied. A *denied request* is a request which is rejected by the airline due to lack of availabililty of space. Space is categorized by fare class on each flight. A *fare class* is a grouping of similar published fares created for the purpose of controlling reservations. Each fare class is assigned a certain number of spaces on each flight. Generally, but not always, the highest fare class contains the highest fares, the next highest fare class contains the next highest fares, and so forth. The lowest fare class is usually comprised of the lowest fares.

When the spaces corresponding to a certain fare class are filled, we say that the fare class is *full* or *closed*. If the corresponding fare class is closed, then the request for travel is denied. The reservations agent usually attempts to accommodate the request in a different fare class on the same flight or in the requested fare class on another flight. If the traveler accepts a different fare class on the same flight, then the airline has made a *vertical recapture* of the traveler. If the traveler is accommodated on another flight in the requested fare class, then the airline has made a *horizontal recapture* of the traveler (Belobaba, 1987). Otherwise, if the traveler opts not to fly or chooses a different airline, the traveler is *lost* to the airline.

Ticketing occurs when a traveler pays the fare associated with the reservation to an airline agent. In exchange, the airline gives the traveler a *ticket*, which verifies the itinerary of the trip and the fare paid. Ticketing is usually done at one of three points during the booking process depending on the rules associated with the fare paid and the policy of the particular airline. Most frequently, a traveler purchases a ticket between the time that the reservation is made and the time that passengers board the aircraft on the departure date. We should note that a cancellation may still occur after the ticket is issued. However, it is also possible that a reservation is cancelled before a ticket is ever purchased. This might happen because of a change in plans or a traveler's non-compliance with the stated ticketing rules of the airline. Finally, in isolated cases, a traveler may purchase a ticket after boarding the aircraft enroute to his destination. The Eastern Airlines Boston-New York-Washington Shuttle used to allow enroute purchases of tickets as did the now-defunct PeoplExpress Airlines.

2.2.2 The Cancellation Phase

The second step of the booking process, the cancellation phase, starts after a reservation is made. In the days and weeks before the flight departs, the traveler may cancel a reservation. A *cancellation* is a reservation which is cancelled (explicitly or implicitly) before a flight departs. An *explicit* cancellation occurs when the traveler cancels a reservation due to a change in travel plans. From the perspective of a particular fare class on an individual flight, any change in plans is reflected as a cancellation, regardless of where the passenger is re-booked. For example, when a traveler cancels a reservations in fare class 2 and re-books in fare class 1 on the same flight, fare class 2 experiences a cancellation and fare class 1 gains an additional reservation.

An *implicit* cancellation occurs when the airline cancels a reservation due to a traveler's non-compliance with the airline's stated terms and restrictions for that reservation. For example, many of the lowest fare reservations must be ticketed within 24 hours after the reservation is made. If the traveler does not purchase the ticket within 24 hours, the airline automatically cancels the reservation. A *no-show* is a last minute cancellation by the traveler. In general, a no-show is a traveler with a reservation who does not show up at the airport on the day of the flight. Each airline has rules regarding the minimum time before departure of a flight by which the traveler must show up. If a traveler does not show up by the minimum time, the reservation may be cancelled. Travelers become no-shows if they do not show up to board the flight.

2.2.3 The Boarding Phase

The boarding phase is the third step of the booking process, which occurs at the airport on the day of the flight. If a reservation is not cancelled before the flight departs, the traveler shows up at the airport and becomes a passenger. A *passenger* is a traveler holding a reservation who shows up at the airport on the day of departure before the designated minimum check-in time. If there are sufficient seats on the aircraft, then the passenger is allowed to board the aircraft. A *passenger boarded* is a passenger who obtains a seat on the aircraft and departs on the scheduled flight.

On the other hand, the flight may be *oversold*. That is, the flight is overbooked and more passengers show up at the airport than the capacity of the aircraft. In this case, some passengers will be denied access to the aircraft. A *denied boarding* is a passenger who is unable to board the aircraft due to a lack of seats. A denied boarding may be *voluntary*, where the passenger volunteers to not board the aircraft in exchange for some type of compensation. Otherwise, the denied boarding is *involuntary*, where the airline refuses to allow the passenger to board the aircraft, and compensation is granted to the passenger according to the policies of the airline.

2.3 The Booking Process: A Microeconomic Framework

As described in the previous section, the booking process consists of a series of interactions between a potential air traveler and an airline. In microeconomic terms, the airline booking process is an economic interaction between a utility-maximizing consumer, the potential air traveler, and a profit-maximizing producer, the airline firm. The consumer decides whether to travel by air and creates flow over an airline's network of flights. Taken together, consumers create the demand for air travel services. The producer is the airline which provides a schedule of flights between city pairs and makes available a certain number of spaces in each fare class on every flight. Together, the world's airlines produce the supply of air transportation.

This section first describes the supply of air transportation and the profit maximization goals of an individual airline. Next, consumer demand for air travel and the utility maximization goals of the consumer are discussed. Third, this section examines the mechanics of how airline bookings are actually made, through an airline agent and on computer reservations systems, and the effect on the interaction between the potential air traveler and the airline. Finally, we discuss the booking data resulting from the economic interaction between the travelers and the airline. In conclusion, we show how the need for forecasting of the booking process arises from this economic interaction.

2.3.1 Supply of Air Transportation: Schedule of Flights

As the producer in the microeconomic context, an airline provides the supply of air transportation. The supply takes the form of a schedule of air services between a set of origins and a set of destinations. Let us define the basic terminology of the air transportation schedule. (Simpson, 1982) A *route map* is a geographical network of air service connecting the cities to be served. A partial route map of a major U.S. airline is given in Figure 2.2. A *market* is an origin-destination pair of cities. In Figure 2.2, Boston-Milwaukee is a market, so is Boston-Seattle. A *link* connects two cities on the route map with an aircraft flown non-stop. Figure 2.2 shows that Boston-Milwaukee is a link and Milwaukee-Seattle is a second link in the network. A *flight leg* is a link flown by an aircraft at a specific departure time. For example, flight 1234 from Boston to Milwaukee at 4:00 p.m. is a flight leg.

A *route* is a consecutive series of flight legs flown by a single aircraft. For instance, Boston-Milwaukee-Seattle is a route made up of two flight legs. A *flight* is a route flown by an aircraft at a specific time. For example, on a Boston-Milwaukee-Seattle route, there might be two flights: flight 65 at 9:00 a.m. and flight 345 at 6:00 p.m. Finally, a *path* is a series of flights used by an air traveler from his origin to his destination. If two or more flights are taken by an air traveler, then we say that the air traveler makes *connections* between flights. To illustrate, if a traveler is going from Philadelphia to Phoenix, his path may be Philadelphia-Milwaukee on flight 185 and Milwaukee-Phoenix on flight 225. In this case, the traveler connects from flight 185 to flight 225 at Milwaukee.

2.3.2 Supply of Air Transportation: Fleet Assignment

After an airline decides the routes that it wishes to fly, the next major aspect of the supply of air transportation is *fleet assignment*. In brief, given that an airline has different types of aircraft in its fleet, it must decide which type of aircraft will fly on each route. For example, Northwest has 12 aircraft types in its fleet. Table 2.1 shows the aircraft types. Each aircraft type has a cruising speed, maximum range, one or more seating configurations, and certain performance characteristics, such as noise levels from the engines and minimum required runway length. The length of the flight, the runway length at the origin and destination airports, and any noise abatement rules determine whether an aircraft can feasibly fly on a particular flight leg. The seating configuration, or number of seats on the aircraft, is crucially important to the profit maximization goals of the airline. Aircraft capacity directly determines how many passengers are able to board a particular flight and, hence, places an upper bound on the potential revenues generated from the flight. Furthermore, the type of aircraft used on a particular flight largely defines the cost of operating the flight. For more detail on the fleet assignment problem, see Simpson (1982).

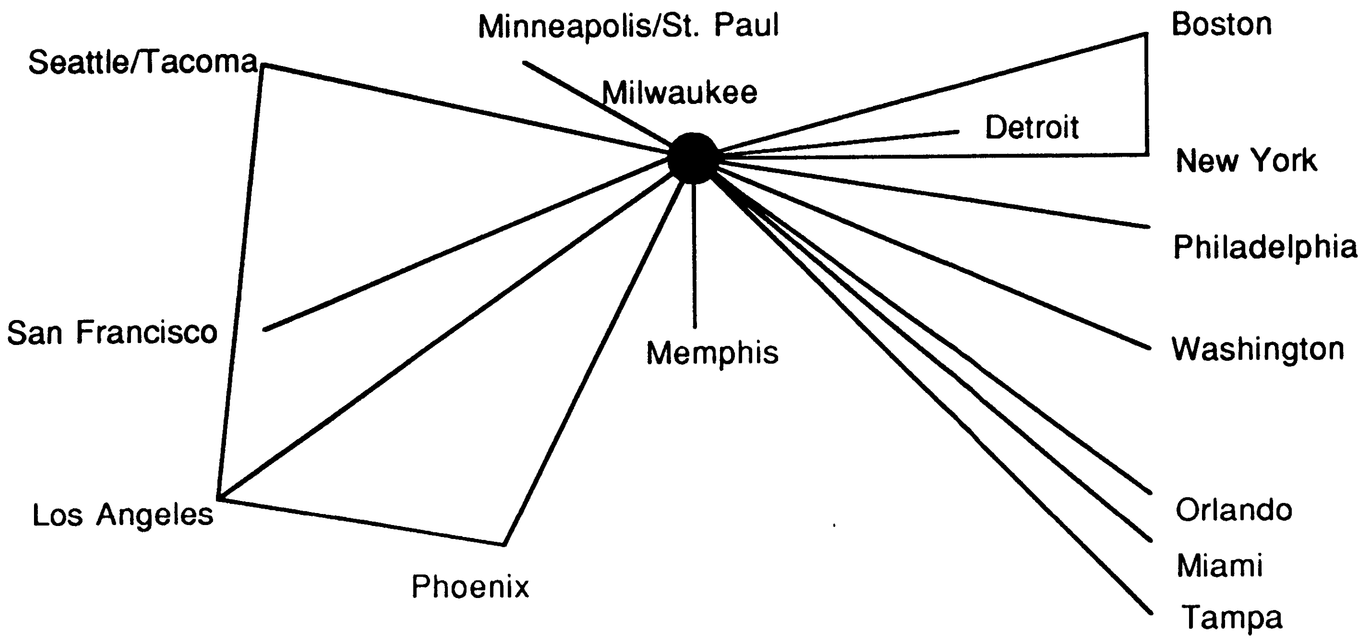

Figure 2.2 Partial Route Map of a Major U.S. Airline

AIRCRAFT TYPE	NUMBER OF SEATS
747–400	420
747	400
DC–10	284
DC–10 (International)	294
757–200	184
A–320	150
727–200	146
MD–80	143
DC9–50	122
727–100	118
DC9–30	100
DC9–10/20	78

Table 2.1 Aircraft Types of a Major U.S. Carrier

2.3.3 Supply of Air Transportation: Aircraft Cabins and Fare Classes

Once the routes of an airline are fixed and the fleet has been assigned, the third major element of the supply of air transportation is how individual seats are sold aboard the aircraft. First, we examine the arrangement of an aircraft cabin. An aircraft is usually physically divided into several cabins, each offering a different level of "on-board service". A typical three cabin aircraft includes a First Class cabin, a Business Class cabin, and an Economy (Coach) Class cabin. The First Class cabin is generally located in the front of the aircraft and offers the highest level of on-board amenities, including more flight attendants per passenger, wide seats, complimentary cocktails and movies, and enhanced meal service. The Business Class cabin typically is located between the First Class cabin and the Economy Class cabin. Business Class offers an intermediate level of in-flight amenities, often with complimentary cocktails and movies as well as better meal service. Finally, the Economy Class cabin is located in the rear of the aircraft and includes a basic level of in-flight amenities. Although a meal is often served, passengers must usually pay for additional amenities, such as cocktails and movies. This three cabin system is common on international routes, as well as on European and Canadian domestic routes. On U. S. domestic flights, most airlines have a two class system with a First Class cabin and an Economy Class cabin.

To complete the description of the supply of air transportation, we must examine the fares charged in the various cabins of the aircraft. In each market, a large number of published fares are available. A *published fare* in a market includes the price of the trip for a given level of service and any accompanying rules, travel restrictions, and effective dates. Airlines group the published fares into *fare classes* for the purpose of controlling bookings in their reservations systems. A fare class is designated by a single letter code, such as "F", "C", or "Y". Currently, the major airlines use from 5 to as many as 14 fare classes for seat inventory control purposes.

Each fare class is assigned to a physical cabin on the aircraft. However, in most cases, there are more fare classes than physical cabins on the aircraft. Thus, travelers booked in two or more different fare classes may sit in the same cabin. Since the Economy Class cabin usually has the largest number of seats and airlines publish a wide variety of Economy fares, the majority of the fare classes are assigned to the Economy Class cabin.

For example, Northwest Airlines currently has 12 fare classes on domestic flights (see Table 2.2). The First Class cabin has two fare classes: F and A. The F fare class remains the full First Class fare designator. The A fare class identifies discounted First Class fares. For example, frequent flyers traveling on free or low cost upgrades in the First Class cabin are booked in the A fare class. Similarly, the Business Class cabin (if available) has two fare classes: C and D. The C fare class remains the full Business Class fare designator. The D fare class has become the designator for discounted Business Class travel.

In the Coach cabin, there are 8 fare classes: Y, B, M, H, Q, V, W and K. The Y fare class corresponds to the full Economy fare. The B fare class is composed of slightly less expensive fares with some minor restrictions. The bottom five fare classes, M, H, Q, V, and, K identify increasingly discounted fares and, usually, increasing service restrictions. These fares all have significant advance purchase, minimum stay, limited change and non-refundability restrictions. Additionally, spaces allocated to these bottom fare classes may be severely restricted or unavailable on some flights. The W fare class is reserved for frequent flyer award tickets.

FARE CLASS CODE	DESCRIPTION	PERCENTAGE OF FULL COACH FARE
F	First Class	150%
A	First Class Discount	varies
C	Business Class	120%
D	Business Class Discount	varies
Y	Full Coach	100%
B	Shallow Discount Coach	75%
M	Moderate Discount Coach	60%
H	Moderately Deep Discount Coach	45%
Q	Deep Discount Coach	35%
V	Very Deep Discount Coach	25%
W	Frequent Flyer Award Coach	varies
K	Virtually Free Coach	10%

Table 2.2 Fare Classes of a Major U.S. Airline

2.3.4 The Airline's Perspective: Profit Maximization

The airlines, like any other industry, seek to maximize profit. Since profit equals revenue minus cost, profit can be increased by decreasing the costs of operating the firm and by increasing the amount of revenue collected. In the years immediately following the U.S. Airline Deregulation Act of 1978, the airline industry attempted to minimize long-run costs by enactment of two-tier labor contracts, acquisition of fuel-efficient aircraft, and improved utilization of employees. More recently, the airlines have become concerned with the shorter run aspects of profit maximization: deciding which aircraft to operate on each flight and how many spaces to allocate to each fare class on each flight.

Since demand is stochastic in nature, the airline's short run goal is to maximize total expected profit (total expected revenue - total cost). First, total expected revenue is a function of fare, quantity of spaces allocated to a fare class, and the probability distribution of selling the spaces. Let the fare on flight f in fare class c departing on date d be F_{cfd}, the quantity of spaces allocated to flight f in fare class c departing on date d be CAP_{cfd}, and the probability of selling q spaces in fare class c on flight f departing on date d be $p_{cfd}(q)$. The airline's total expected revenue is the product of the fares paid and the expected number of spaces sold, summed over all classes, flights, and days of the year. Mathematically,

$$TER = \sum_{c=1}^{C} \sum_{f=1}^{F} \sum_{d=1}^{D} (F_{cfd} * \sum_{q=0}^{CAP_{cfd}} p_{cfd}(q)*q)$$

where TER is the total expected revenue of the airline. Note that the probability distribution of selling q spaces is censored at CAP_{cfd}. That is,

$$p_{cfd}(CAP_{cfd}) = \sum_{q=CAP_{cfd}}^{\infty} p_{cfd}(q).$$

The issue of censored demand is discussed at length in Chapter 5.

Since the deregulated air travel environment is generally controlled by several large airlines, it is an oligopoly from the microeconomic standpoint. One primary characteristic of many oligopolistic industries is a "kinked" demand curve, which is elastic at points above the equilibrium price and inelastic at points below the equilibrium price. (Baumol and Blinder, 1979) That is, if a firm raises prices, it faces an elastic demand curve and would observe a large loss of demand to its competitors. However, if a firm lowers its prices, it faces an inelastic demand curve. All competing firms would match the lower price and the increase in demand would be comparatively small. Thus, in the airline industry, there is little incentive to change air fares in the short run. Recent experience in the U.S. verifies the relative short run stability of air fares. (Power, 1990) As a result, we assume that the price F_{cfd} of travel is fixed in the short run. Therefore, the short run revenue decision variable for the airline is CAP_{cfd}, the number of spaces allocated to fare class c on flight f on date d in order to maximize total expected profit.

Two main constraints arise for the airline on the revenue side of the profit maximization process. First, we assume that the schedule of flights is fixed in the short run. Therefore, the capacity on each flight is limited to the total authorized capacity of aircraft a on flight f departing on date d, $TCAP_{afd}$. The second constraint is that the demand on each flight is probabilistic in nature. An airline never knows with certainty how many requests will be received for a particular flight. We must assume a specific probability distribution of demand (a specified form of $p_{cfd}(q)$) in order to calculate the optimal number of seats to sell at each fare.

Total cost is determined largely by the length of the flight and the type of aircraft used on a given flight. Let k_{afd} be the cost of using aircraft a on flight f departing on date d and z_{afd} be a binary variable which equals 1 if aircraft a is used on flight f departing on date d and 0 otherwise. Then,

$$TC = \sum_{a=1}^{A} \sum_{f=1}^{F} \sum_{d=1}^{D} k_{afd}{}^{*}z_{afd}$$

where TC is the total short run cost of the airline.

Several constraints arise on the cost side of the profit maximization process. First, it should be noted that aircraft fly on a network and must be assigned on a network. For instance, if an aircraft arrives at an airport, it also must leave the same airport at a later time. Thus, the fleet assignment must be balanced. In the mathematical formulation, we will call these constraints the *network constraints*. Second, there are airport considerations such as maximum available gate space and minimum turnaround time (time between landing and subsequent takeoff). In the mathematical formulation, these constraints are called the *airport constraints*. Third, we must assign an aircraft to routes within its performance capabilities. For example, an airline must consider the aircraft's maximum range and the runway length at the origin and destination airports. If a certain aircraft a is unable to fly flight f, we set the cost of operating aircraft a on flight f for all dates d, k_{afd}, to infinity (or a very large cost).

Mathematically, the airline's profit maximization problem can be stated as follows:

$$\text{Maximize} \quad TER - TC = \sum_{c=1}^{C} \sum_{f=1}^{F} \sum_{d=1}^{D} (F_{cfd} \, {}^{*} \sum_{q=0}^{CAP_{cfd}} p_{cfd}(q)^{*}q) - \sum_{a=1}^{A} \sum_{f=1}^{F} \sum_{d=1}^{D} k_{afd}{}^{*}z_{afd}$$

subject to

$$\sum_{c=1}^{C} CAP_{cfd} \leq \sum_{a=1}^{A} TCAP_{afd}{}^{*} \, z_{afd} \text{ for all flights f and dates d}$$

NETWORK CONSTRAINTS

AIRPORT CONSTRAINTS

z_{afd} = 0 or 1 for all aircraft a, flights f, and dates d

$CAP_{cfd} \geq 0$ and integer for all classes c, flights f, and dates d

Thus, the airline wishes to maximize expected profit by deciding which type of aircraft to assign to each flight and how many seats to sell in each fare class on every flight. However, the problem is constrained by the network constraints, airport characteristics, aircraft capacity, and the probabilistic nature of demand.

In the very short run, we assume that the airline has already made the fleet assignment for each route. In effect, the z variables are fixed in the above problem. Thus, in the very short run, the airline simply wants to determine how many spaces to allocate to each fare class. We call this problem the *revenue maximization* problem. It is formulated as follows:

$$\text{Maximize} \quad TER = \sum_{c=1}^{C} \sum_{f=1}^{F} \sum_{d=1}^{D} (F_{cfd} * \sum_{q=0}^{CAP_{cfd}} p_{cfd}(q)^* q)$$

subject to

$$\sum_{c=1}^{C} CAP_{cfd} \leq TCAP_{fd} \text{ for all flights f and dates d}$$

$CAP_{cfd} \geq 0$ and integer for all classes c, flights f, and dates d

where $TCAP_{fd}$ is the capacity of the aircraft assigned to flight f departing on date d. We note that the network constraints and the airport constraints disappear from the formulation because they are solely related to fleet assignment. In this thesis, we focus primarily on the very short run revenue maximization problem.

2.3.5 The Consumer's Perspective: Utility Maximization

Classical microeconomic theory assumes that individuals will make choices which are the most favorable to them. The measure of favorability or goodness of a particular alternative is called utility. Thus, rational consumers attempt to maximize their utility. Applying this concept to air travel, a consumer desires to maximize his or her utility function when making a request for travel. The main factors involved in making travel plans are: travel dates, price, service, and restrictions (Belobaba, 1987). From an economic standpoint, the traveler usually has particular travel dates under consideration when calling an airline agent. Denote the desired travel dates by d^*. In addition, we expect that a traveler has a certain "target" price that they are willing to pay. Let us call the desired price p^*.

Third, a potential traveler in general has an expected level of service. The level of service includes both on-board components as well as routing components. An example of the on-board component is when the consumer desires Business Class service on a particular trip. The routing component measures the quality of the itinerary in terms of the routing of the trip. For example, a business traveler who is time-conscious may prefer a non-stop flight over connecting or multi-stop flights. Denote the desired level of service by s^*. Finally, when requesting a flight, a consumer knows the flexibility of the trip. That is, he knows what type of restrictions, if any, he is willing to accept on his trip. For example, a business traveler may opt for a fare which can be changed at the last minute. However, on a leisure trip, a consumer may be willing to accept a lower, non-refundable and non-changeable fare. Let us denote the desired set of restrictions by r^*.

A reservation is an economic bundle characterized by (d,p,s,r). As discussed above, the consumer's desired bundle, or ideal point, is (d^*,p^*,s^*,r^*). A consumer's utility can be measured as a function of how a particular reservation (d,p,s,r) differs from the consumer's ideal point (d^*, p^*, s^*, r^*). Mathematically,

$$U = h(d - d^*, p - p^*, s - s^*, r - r^*)$$

where U represents the utility function of an individual and h is a general function. So, to maximize the overall "desirability" of travel, the consumer maximizes U over all possible reservation bundles. The set of all possible reservation bundles (d,p,s,r) is called the *choice set* in economic theory.

The choice set in air transportation has four important characteristics. First, as illustrated above, the choice set is *multidimensional.* A choice of air travel is defined by the travel date, price, level of service, and restrictions. In addition, the choice set is *bounded* on the travel date and price dimensions. Most air travelers have constraints on departure dates and return dates. For example, a business traveler often has to leave on one particular date and return on another specific date. A leisure traveler may have a limit on the length of a vacation and a specific time frame for the trip, such as a particular month. On the price dimension, there is often an upper bound on the price of the trip. Leisure travelers are usually the most price sensitive, facing underlying budget constraints.

A third important aspect of the choice set of reservation bundles is that the air traveler often does not see the entire feasible set of reservation bundles. Most requests are made through an airline reservations agent or a travel agent. The traveler often relies on the agent to reduce the choice set to the two or three "most appropriate" choices. Therefore, the actual choice set for a particular traveler may be quite *limited.* Finally, the airline booking process is dynamic. Many additional reservations are made each day interspersed with cancellations of existing reservations. The airlines are constantly revising the availability of seats in each fare class. Also, restrictions on the lowest fares limit their availability as the day of the flight approaches. Therefore, the available choice set *changes over time*. In general, the closer the departure date, the smaller the choice set of reservations bundles.

In mathematical terms, the statement of the consumer's utility maximization problem is the following:

$$\text{Maximize } U = h(d - d^*, p - p^*, s - s^*, r - r^*)$$

$$\text{subject to } (d,p,s,r) \in \mathbf{C}(d_e, d_l, p_{max}, ag, t)$$

where **C** denotes the choice set which is a function of d_e, the earliest possible departure time, d_l, the latest possible departure time, p_{max}, the maximum fare that the traveler is willing to pay, ag, the type of airline agent used (determines the number of choices given to the traveler), and t, the time (number of days) before departure at which the reservation is made. In conclusion, the consumer wishes to maximize utility by finding a reservation bundle very close to the desired bundle. However, the choice set is limited due to bounds on travel dates and price, unavailability of information about all possible reservation bundles, and the amount of time before departure.

2.3.6 How Bookings Are Made: The Role of Travel Agents and Computer Reservations Systems

The transaction between the airline and the consumer takes place through an agent of the airline and is recorded in a computer reservations system (CRS). This section examines the effect of the airline agent and the computer reservations system on the booking process.

An economic agent is a person hired to represent the airline. The principal type of agent for the U. S. air transportation industry is the travel agent. Travel agencies are independent companies which distribute tickets for airlines, prepare travel itineraries, reserve hotel rooms, and arrange many other travel services for the traveler. The airlines pay a commission to the travel agencies for each ticket sold. As of 1984, travel agencies sold over 70% of all airline tickets. (Davis, 1987) In contrast, the airline's own sales representatives handle only approximately 25% of the total airline tickets. Other minor airline agents include tour operators, who coordinate group tours and sometimes deal directly with the consumer, and wholesalers, who in some cases provide the link between the airline and the consumer (Simpson, 1982). By far, travel agents are the dominant force in selling airline tickets.

Travel agents can affect the booking process in several key ways. First, the profitability of a travel agency is dependent on the satisfaction of the consumer. Therefore, a travel agent has an incentive to try to meet the exact needs of each potential air traveler. If a traveler is a regular customer of a travel agency, the travel agent may know the traveler's particular travel requirements. Thus, the travel agent may be able to match the air traveler's utility function more closely than an airline reservations agent. In this case, the travel agency may be aiding the airline and the consumer by closely matching the airline's schedule of flights and the consumer's travel needs.

Second, a travel agent may affect the booking process by reducing the consumer's choice set. Rather than presenting a consumer with a large number of alternative itineraries, a travel agent may limit the choice set to two or three alternatives. In addition, an airline may give travel agencies extra compensation, called *override commissions*, for booking large numbers of travelers on its flights. Thus, the travel agent may try to book as many travelers as possible on this airline, which directly affects the booking process. For example, suppose airline A and airline B have equivalent non-stop service between Boston and Washington, D. C. If airline A gives local travel agencies incentives for large numbers of bookings, airline B (if it does not immediately match the incentives) may experience a large drop in bookings on its flights in the market.

Third, large travel agencies can sometimes obtain exceptions to the booking limits on fully booked fare classes. In essence, the airline allows the travel agent to book a space in a fare class which is closed. This affects the booking process by increasing the possibility of oversales (and, hence, denied boarding compensation) on fully booked flights and directly thwarts the airline's attempts to maximize revenue. Alternatively, if the desired fare class is closed and an override is not possible, a travel agent may knowingly issue a ticket for a flight without making a reservation. In this case, the traveler is called a "*no-rec*", since there is no record of the reservation. No-recs can affect the booking process by increasing the possibility of an oversold flight.

The computer reservations system (CRS) is the computerized system on which reservations are recorded. During the 1960's, all of the major U. S. airlines purchased in-house computerized reservations systems. The CRS allowed the airlines "... to match passengers to seats, speed communications among all airlines, contain seat availability on carriers' schedules, and put terminal access in the office of travel agents." (C. R. Smith from Davis, 1987) After several collective attempts to develop a common CRS for travel agencies during the late 1960's and early 1970's, American Airlines, United Airlines, and TWA individually made heavy investments and developed CRS's for travel agencies (and, more recently, consumers on home personal computers). Later, Texas Air and Delta put computer reservations systems in travel agencies. As of 1987, American's SABRE system and United's APOLLO system were by far the most popular systems among U. S. travel agencies. Together, the two systems were in over two-thirds of the travel agencies in the U. S.

Since two major airlines, American and United Airlines, have control of the major computer reservations systems, there is an inherent conflict of one airline providing information on its competitors' flights. In fact, a number of biases have been found which directly impact the booking process. (Davis, 1987) First, host airlines (owners of the computer reservations systems) can bias the computer screen in order to give their own flights preferential display. This may influence the decision of the travel agent toward the host airline's flights and could result in a decrease in bookings on a competitor's flight. As a result of complaints by smaller airlines, the U. S. government has introduced a set of rules which have removed much of the bias from the CRS. Therefore, the impact of these biases has decreased considerably, although some airlines remain concerned about "subtle" biases in the CRS. (Taib, 1990)

Second, host airlines may have access to their competitors' booking data. The host airlines could conceivably use this data to make supply, capacity, and pricing decisions in order to alter the travelers' booking patterns. However, it appears that the vast amount of data has dissuaded any host airline from effectively using competitors' booking data. A third effect of the CRS is that host airlines can have real time communications between the CRS and their in-house seat inventory control system in order to continuously update availability of spaces and fares. Other airlines may have to wait several hours or overnight to update availability on their flights. When spaces are selling quickly, the host airline has an obvious advantage over the other airlines.

Finally, the CRS has allowed airlines to accurately collect booking data, keep track of current bookings on flights which have not yet departed, and impose booking limits on fare classes. In general, the CRS has increased the accuracy and complexity of the transactions between the airline and the traveler.

2.3.7 The Result of the Economic Interaction: Booking Data

The economic interactions between potential travelers and the airline result in reservations and cancellations in each class on each flight. We define *total bookings* at any time before departure as the total reservations made up to that time minus the total cancellations made. That is, bookings are the total number of reservations remaining (which have not been cancelled) in the system at any time prior to departure. The *booking curve* is the graph of total bookings versus the number of days before departure. A sample booking curve is presented in Figure 2.3. In this section, we examine the characteristics of the booking curve and, then, describe the booking data available to the airline.

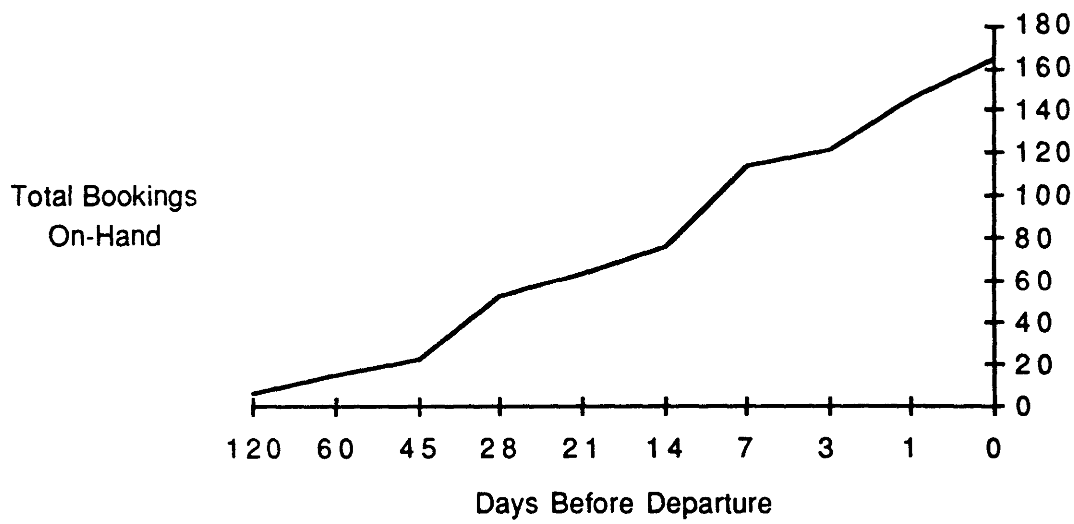

Figure 2.3 Sample Booking Curve

The characteristics, or shape, of the booking curve of a specific class on a particular flight are determined by four key factors. The first factor is how early (prior to departure) reservations occur. The time at which reservations are made depends on the type of market and the fare class of interest. While leisure markets tend to book early, business markets tend to book late. Similarly, low fare classes which appeal to leisure travelers usually book early. Higher fare classes, used by business travelers, generally book late. (See Figure 2.4.)

The second important factor in determining the shape of the booking curve is how fast reservations are made in any period prior to departure. The slope of the booking curve measures the rate of bookings in a given period. The rate of bookings depends on the fare class, the fare levels, and any promotions such as increased advertising. Lower fare classes tend to have high booking rates just before advance purchase restrictions prevent further reservations. Higher fare classes generally exhibit high booking rates within a few days of departure date. (See Figure 2.4.) Furthermore, fare levels and promotions may affect the rate of bookings. Relatively high fares may slow the rate of bookings; relatively low fares or heavy promotion of a particular destination may accelerate the booking rate.

The third factor is the effect of booking limits. Each class on every flight is constrained by a maximum authorized booking limit. Once the booking limit is reached, no additional reservations are accepted. Thus, the booking curve "flattens out" at the booking limit. An example of this phenomenon is given in Figure 2.5. Since the airline cannot observe the demand for reservations greater than the booking limit, the airline sees only constrained booking data on high demand flights. In effect, the airline loses information about the true, underlying booking curve when the booking limit is reached.

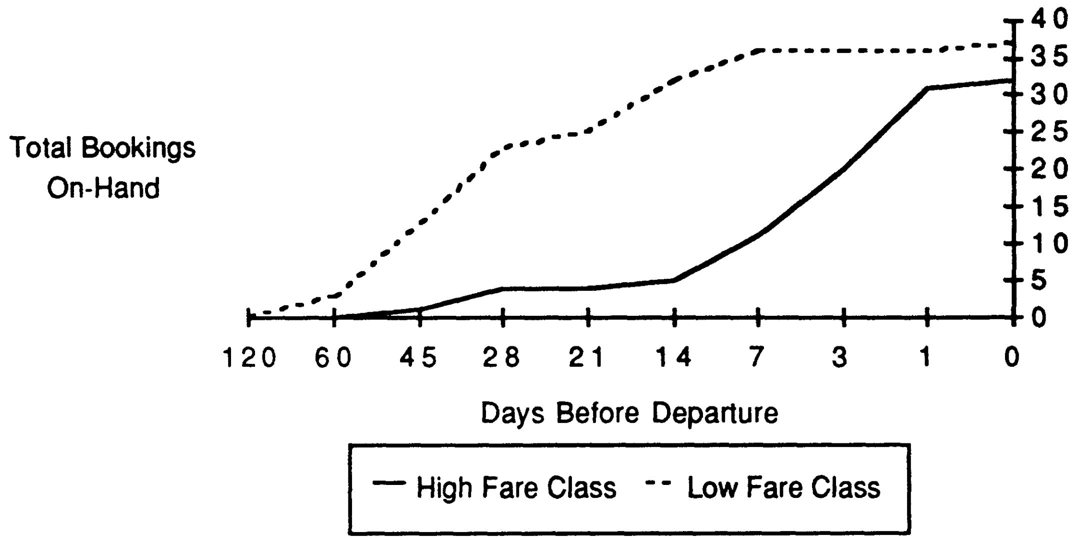

Figure 2.4 Sample Booking Curves of High Fare Class and Low Fare Class

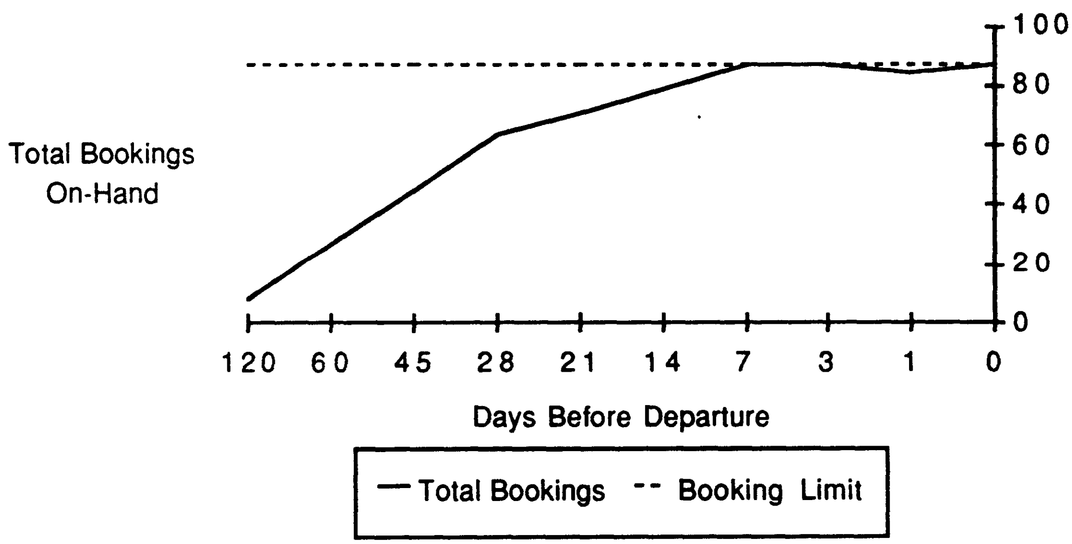

Figure 2.5 Sample Booking Curve Limited by a Booking Limit

Finally, the booking curve is affected by the time at which cancellations occur during the booking process. If cancellations on any day exceed additional reservations made, a downward turn is observed in the booking curve. An example of this condition is displayed in Figure 2.6. The occurrence of cancellations depends on the type of market and the fare class. In a few pure leisure markets, the airlines allow travel agencies to make reservations without actually purchasing a ticket. As a result, many cancellations may occur prior to departure - - if the travel agents are unable to sell all of their reservations.

Fare rules for bookings made in higher fare classes usually allow a passenger to change their itineraries without penalty. Thus, high cancellation rates and no-show rates are often observed in the high fare classes. In the lower fare classes, the non-refundability restrictions strongly discourage cancellations of ticketed reservations. Many of the lower fare class reservations are automatically cancelled if not ticketed within 24 hours. Hence, an airline may observe high cancellation rates in the lower fare classes because of the automatic cancellation provision. In summary, four major factors affect the characteristics of the booking curve of a particular class on given flight: how early reservations occur, how fast reservations arrive, the effect of the booking limits, and when cancellations occur during the booking process.

For the airline, the booking curve is the primary source of booking data. Most airlines store at least several months of historical booking curves for each class on flights which have already departed. Additionally, the airlines have partial booking curves up to the current day for each class on flights which have not yet departed. In most current seat inventory control systems, the booking curves do not contain information about reservations and cancellations separately. Instead, the "net" booking curve only identifies the total number of bookings in a fare class on a particular flight. Most airlines could obtain reservations and cancellations separately, but the cost of storing the additional data may be substantial. Hence, at any time

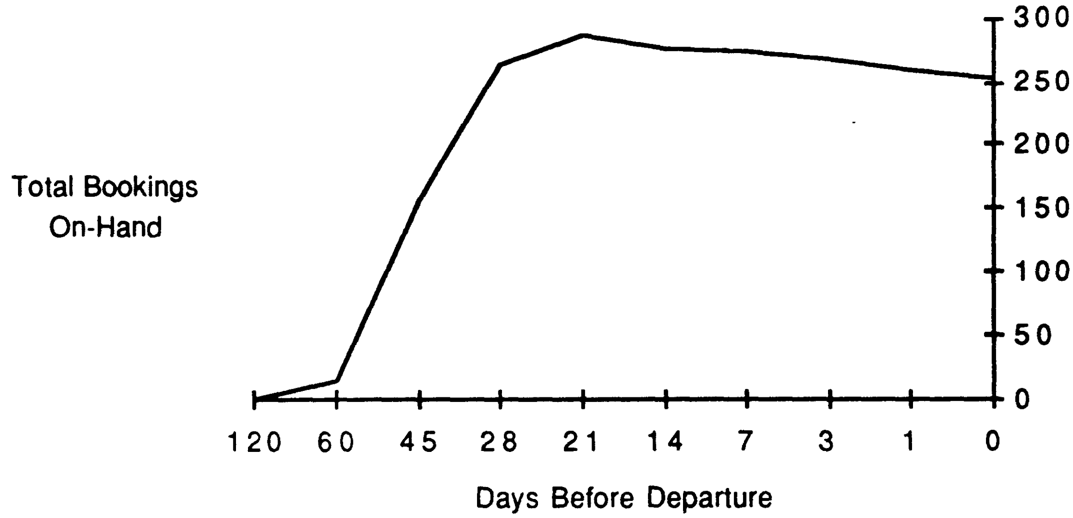

Figure 2.6 Booking Curve with a Decrease in Bookings

before the departure of a flight, the booking data available is the total number of bookings at each point on the partial booking curves and the completed historical booking curves from previous departures of the same flight number.

2.3.8 The Need for Forecasting the Booking Process

The utility maximization goal of the consumer and the profit maximization goal of the airline firm, as well as the role of the travel agent and the CRS, are inherently intertwined. The consumer sees a choice set which is influenced by the airline's profit maximization decisions, the travel agent, and the CRS. On the other hand, the airline faces uncertain demand for its seats which depends heavily on the utility maximization decisions of the consumer. It is exactly this latter process which we will model in this thesis.

Given that the airline wants to maximize profit, it requires an accurate forecast of total bookings in each fare class. Because it places limits on the number of seats sold in each fare class, the airline only views constrained booking data. In particular, the airline sees only the accepted demand and does not observe the demand that was turned away because of the booking limits and the finite capacity of the aircraft. In making future revenue maximization decisions, the airline wants to know the total underlying bookings -- not simply the observed bookings. In this thesis, we will develop statistical and probabilistic techniques to use observed data from booking curves and booking limits to forecast the total underlying demand. Furthermore, the analysis in Appendix A indicates that more precise knowledge of the actual underlying demand can aid the airline significantly in making optimal revenue maximizing decisions.

Chapter 3 Previous Approaches: A Literature Review

3.1 Introduction

The previous chapter defined the booking process in detail, presented a microeconomic framework for the airline booking process, and concluded with the need for forecasting on a class by class basis on each flight departure. In this chapter, previous approaches to airline reservations forecasting are explored. First, we categorize the different types of forecasts used in the airline industry. The three primary categories are macro-level, choice modeling, and micro-level. We define each of the categories and discuss their relevance to the forecasting problem at hand. Then, we describe the relevant literature on statistical modeling for each of the three categories. In the third section of the chapter, we discuss the relevant probabilistic literature for micro-level forecasting. The final section of the chapter explores the approach taken in this thesis.

3.2 Types of Forecasts in the Airline Industry

The types of forecasts in the airline industry fall into three primary categories: macro-level, choice modeling, and micro-level. *Macro-level* forecasts involve large-scale, aggregate forecasts of total airline passengers. Examples of macro-level forecasts include projections of total annual domestic air travel and forecasting future U.S. to Europe air travel over the next decade. *Passenger choice modeling,* on the other hand, is the study of a single individual's behavior based on socioeconomic factors and the characteristics of alternative travel options. Illustrations of passenger choice modeling are predicting an individual's choice of air travel as opposed to rail travel or forecasting the choice of a particular airline over alternate airlines.

The final type of forecasting in the airline industry is *micro-level* forecasting, which is the primary focus of this thesis. Micro-level forecasting includes forecasting passenger demand by flight, date, and fare class. In terms of level of aggregation, micro-level forecasting falls between macro-level forecasts and passenger choice modeling. An example of micro-level forecasting is the prediction of total demand in the Y fare class on flight 1234 departing on September 22, 1990 at a point 14 days before departure.

3.2.1 Macro-level Forecasting Literature

The principle references in macro-level forecasting are Taneja (1978) and Kanafani (1983). Taneja devotes an entire book to regression models for aggregate airline traffic forecasting. He addresses statistical methods for macro-issues such as forecasting total airline traffic and projections of future national traffic growth. Taneja covers regression analysis in great detail including multiple regression models, tests of model assumptions, model specification, model selection, and multiequation regression analysis. The principal focus in the examples and methodology is on aggregate forecasting.

Kanafani (1983) devotes Chapter 9 of his book to the demand for air transportation. He discusses aggregate measures of air travel activity such as passenger volume, aircraft operations, and revenue passenger miles. Kanafani points out that the above measures can be stratified in one of several ways: trip purpose, temporal stratification, origin-destination, length of haul, and type of service (airline, charter, and commuter aviation). The possibility of forecasting by fare type is briefly addressed. However, it focuses on aggregate forecasting for certain fare types. The majority of the chapter on demand for air transportation is devoted to issues in macro-level forecasting, whether it be total U.S. domestic traffic or total annual traffic between San Francisco and Los Angeles.

3.2.2 Passenger Choice Modeling Literature

Passenger choice models in the air transportation are covered briefly by Kanafani (1983). He devotes a section of the chapter on demand in air transportation to passenger choice models. Passenger choice models arise when a traveler must choose among competing services. Kanafani categorizes the types of choices which occur in air transportation, which include route choice, airport choice, airline choice, and fare-type choice. A multinomial logit model is introduced as a method of estimation of passenger choice models. A more general reference to discrete choice modeling in transportation is Ben-Akiva and Lerman (1985).

3.2.3 Micro-level Forecasting Literature

In general, very little research on micro-level forecasting by flight number, fare class, and date has been done. In the 1970's, Littlewood (1972) and Scandinavian Airlines (1978) described some basic characteristics of the airline booking process. In both papers, the authors propose a simple forecasting model for total bookings on a flight based on computing the mean of historical bookings on previous departures of the same flight. The focus in this literature is on forecasting total demand on an entire flight. However, Littlewood and Scandinavian Airlines mention that the models could be used to forecast demand by fare class. The Scandinavian Airlines paper briefly discusses the amount of historical data needed to produce accurate forecasts and the necessity of removing data points corresponding to unusual, non-recurrent events. Overall, the forecasting methodology proposed in these two papers is overly simple, using rather naive statistical methods.

On the academic research side, there are three articles of relevance for micro-level forecasting. First, Sa (1987) presents a rudimentary data analysis based on time series models and regression models. Two ARIMA time series models were estimated for a single fare class on a single flight number. The results of the two estimations were not very encouraging and Sa saw no further merit in time series models. Then, he developed a regression model of the airline booking process. The dependent variable was bookings to come. The explanatory variables included bookings on-hand, a seasonal index, a day of week index, and a historical average of bookings to come. The regression model gave much more positive results. While the results of the estimations show that the regression models have definite merit, Sa did not test the forecasting ability of the models. Furthermore, he did not incorporate the effect of booking limits into the statistical analysis.

The second relevant study is an analysis performed by Brummer et. al. (1988). This study explicitly takes into account that the booking data is constrained by the maximum authorized booking limit. The goal of the study was to find the mean and standard deviation of the true, unconstrained Log-normal distribution, given a data set with some constrained observations. The majority of the study was spent on the mathematical derivation of the likelihood function of a censored Log-normal distribution. This study briefly explored the nature of airline booking data and examined only total bookings on each flight. Brummer et. al. did not perform a class by class analysis of bookings. In addition, there was no attempt to forecast future demand or validate the proposed model on a different data set.

The most relevant research performed on micro-level forecasting is by Ben-Akiva et. al. (1987). This paper proposes three models for flight-specific, class-specific reservations forecasting: a regression model for advance bookings on a given flight, a time series model for historical bookings on previous departures of the same flight number, and a combined advance bookings/historical bookings model. Ben-Akiva et. al. present preliminary analysis using monthly airline data by flight and fare class. The results of the estimations showed that the combined model outperforms the historical bookings and advance bookings models. Although the results indicate potential practical success, the authors did not have sufficient data to validate the results of the estimated models on future flights. Furthermore, the data is monthly, not daily as required in micro-level forecasting of the booking process. Finally, the effect of booking limits was not taken into account in this paper.

As briefly discussed by Brummer et. al., the booking limits placed on each fare class bring about the need to use special methods for estimation of a model with censored data. Maddala (1983) contains a chapter on censored and truncated regression models. He discusses how to estimate censored and truncated models with Normally distributed data, the properties of the estimators, and extensions to other distributions such as Log-normal and Exponential. Schneider (1986) devotes an entire book to truncated and censored samples from Normal distributions. He discusses how to estimate the parameters of the true, underlying distribution, given censored or truncated data. The book contains a set of computer programs which estimate the parameters of various types censored distributions.

3.3 Literature on Stochastic Processes in Reservations Forecasting

The literature on stochastic processes in reservations forecasting is quite limited. In fact, the few authors who have attempted to develop rigorous stochastic models of the airline booking process focus primarily on overbooking. Taylor (1962) develops a stochastic model of the cancellation process. His goal is to find the optimal booking limit such that the aircraft is full on the day of departure. The reservation process is not explicitly considered here. However, he does allow bulk departures of reservations, corresponding to entire groups cancelling simultaneously. The result of this analysis is a complex mathematical probability statement of the current number of bookings remaining in the system. No attempt is made to simplify the probability statement. Because of the complex mathematical probability statement and the state of computer science in 1961, the methodology was not tested on actual airline data.

Rothstein (1971) develops a non-homogeneous Markovian model of the booking process, which requires specification of the distribution of passenger requests and cancellations. However, the goal of Rothstein's paper is to determine optimal levels of overbooking, not to forecast total bookings in a fare class on a specific flight. The Markovian model treats reservations and cancellations separately, rather than as an interspersed process. Additionally, the author (Rothstein, 1985) states that it is not clear whether this probabilistic overbooking model has ever been implemented.

A general reference on stochastic processes is Bailey (1964). Bailey stresses applications to the natural sciences, particularly biological populations. Many population processes (variations of birth, death, and immigration processes) are described in the context of stochastic processes. Bailey also treats the difficult case of time-dependent immigration and death rates. This is crucial for the airline applications, since reservation rates and cancellation rates vary over the duration of the booking process of each flight. If the airline booking process for a fare class on a given flight is considered as a population, then the stochastic models discussed in Bailey's book can be quite relevant to the airline booking process. As in most books on stochastic processes, very little attention is given to parameter estimation and forecasting.

3.4 Current Approach Developed In this Research

The current approach taken in this thesis develops a comprehensive mathematical framework for the analysis of the airline booking process. First, we create a probabilistic model of the airline booking process. This approach, which is the topic of Chapter 4, uses Rothstein's (1971) work as a starting point. In this thesis, unlike Rothstein's work, we consider a stochastic process with interspersed reservations and cancellations. Reservations for a seat on the aircraft are immigrants to the population of travelers in a particular fare class on a specific flight. Cancellations of reservations are similar to deaths of members of the population. We derive a complex probability statement which describes the number of bookings in the system at any time before departure. Then, we introduce a straightforward approximation that simplifies the probability statement considerably. The result is a censored Poisson model of the airline booking process.

Second, we develop a rigorous statistical framework. This approach is the topic of Chapter 5 and uses the research of Ben-Akiva et. al. (1987) as a starting point. We generalize the advance bookings, historical bookings, and combined models. A completely new model, the full information combined model, is formulated and described. In addition, the statistical framework introduces the concept of booking limits on the observed data. To account for the presence of the booking limits, we incorporate the methodology of Maddala (1983) and Schneider (1986) and propose a truncated-censored model. Finally, the statistical framework considers possible alternative demand distributions.

Chapter 6 treats the practical issues in estimation and forecasting of the airline booking process, while Chapter 7 validates the forecasting ability of the censored Poisson model from Chapter 4 and a censored combined model from Chapter 5 on actual airline data. Unlike much of the previous work on reservations forecasting, this thesis tests the forecasting performance of the proposed models. We show that these models not only fit the data well, but also accurately predict demand on future flights.

Chapter 4 A Probabilistic Model of the Booking Process

4.1 Introduction

This chapter analyzes the booking process from a probabilistic perspective. First, the assumption of stationary and distinct demand is discussed. Next, the booking process is described as a stochastic process. Third, with some simplifying assumptions, we introduce the booking process as an immigration and death process. Then, we discuss estimation and forecasting in the context of the probabilistic model. Finally, we describe qualitatively the "ideal" stochastic model, incorporating many of the subtle aspects of the booking process.

4.2 Assumption of Stationary and Distinct Demand

In the next two chapters, we make a fundamental assumption that demand for each fare class is distinct and stationary. The consumer, being presented with a choice set of alternatives as outlined in Chapter 2, is assumed to request one particular fare class which maximizes his/her utility. As a result, the airline is able to distinctly identify the demand for each fare class. We also assume that the demand for each fare class is stationary. In essence, stationarity implies that the data is stable, varying only in systematic patterns. We will examine both of the assumptions in turn.

First, the issue of distinct demand arises because most airlines categorize a number of different fares in a single fare class (as described in Chapter 2). If fares, restrictions, and fare classes stay relatively constant, we would assume that each fare class would identify a distinct grouping of demand. However, in the deregulated environment of the U.S., this is simply not the case. In particular, we describe two key problems with the distinct demand assumption.

The first problem is that, as fares, restrictions, and fare classes change over time, demand in a particular fare class today may not be the same type of demand as last year. For example, V class today may contain non-changeable, non-refundable fares, appealing primarily to leisure travelers. Last year, however, V class may have contained some fares which allowed changes up to the day of departure, appealing to many business travelers. Therefore, over the one year horizon, V class demand is not a distinct, homogeneous grouping of demand. As a result, it is necessary to carefully screen the data in order to guarantee a relative degree of homogeneity over time.

The second problem with the distinct demand assumption is that fares which appeal to the same type of travelers are often found in two or more fare classes. As the number of fare classes has increased in recent years, airlines have categorized fares with similar restrictions in more than one fare class. For example, the non-refundable, non-changeable fares appealing to leisure travelers are often categorized in at least two fare classes. Therefore, one fare class may not distinctly identify a certain type of demand. However, over short periods of time, with few changes in fares, restrictions, and fare classes, an assumption of distinct demand still seems reasonable.

The second assumption is the stationarity of demand over time. Stationarity implies that the data is stable, varying only in systematic patterns. Because of seasonal variations, day of week fluctuations, fare changes, and operational problems, airline demand data is usually not stationary. Pindyck and Rubinfeld (1981) describe a traditional time series decomposition by which we can form stationary data from airline data. Time series data can be represented as the product of four components:

$$B_d = L * S * C * I$$

where B_d are the total bookings on the flight departing on date d, L is the long-term trend in the data, S is the seasonal component, C is the cyclical component, and I is the irregular component (the error term).

The idea is to eliminate the seasonal and cyclical components from the data. Then, the remaining long-term trends and error term represent stationary data, which can be explained by the probabilistic models presented later in this chapter and the statistical models presented in Chapter 5. The primary cyclical fluctuation in airline data is the day of week pattern. Generally, demand levels differ by day of week. Thus, if we use data from the same day of week, then we eliminate the cyclical component from the data. To eliminate the seasonal component from the booking data, the most natural idea is to develop a seasonal index. We can use the seasonal index to deseasonalize the data, thereby removing the seasonal variation. Finally, we should screen the data carefully to delete any severe outliers caused by other factors such as operational problems. By eliminating seasonal variations, day of week fluctuations, and any severe outliers, our data set should be relatively stationary over short periods of time.

In this analysis, we will confine the historical data used for modeling to the past several months (as opposed to the years of data used in similar econometric studies, such as forecasting stock prices, gross national product, etc.) to ensure a homogeneous period of time. Furthermore, all demand analysis will be confined to data from the same flight on the same day of week. In most cases, some type of seasonal adjustment will be made to remove the seasonal effect and screening will be done to eliminate outliers. After the above adjustments are made, the assumption of distinct and stationary demand for each fare class is reasonable.

4.3 The Booking Process as a Stochastic Process

The booking process can be viewed as a stochastic process. The number of bookings in the reservations system for each fare class on a particular flight at any time during the booking process can be characterized as a *random variable*. The *state space*, which is the set of all possible values of the number of bookings in the reservations system, consists of the non-negative integers, 0, 1, 2, 3, ... For revenue maximization purposes, the airlines place limits on the number of bookings allowed in each fare class. If a booking limit is imposed on a fare class, then the state space is reduced to the set of non-negative integers less than or equal to the booking limit. Since the number of bookings is a discrete value (non-negative integer) and the time before departure is a continuous measure, the booking process can be described as a *discrete* state, *continuous* time stochastic process. The goal of this probabilistic analysis is, at any point in time during the booking process, to specify the probability distribution of the number of bookings in the reservations system. Therefore, on future flights for which the booking process is not yet complete, we can use the expected value of the probability distribution to forecast the total number of bookings in each fare class on the day of departure.

4.4 The Immigration and Death Process

As noted in Chapter 2, the airline booking process is characterized by requests for reservations interspersed with cancellations in the weeks and months before a particular flight departs. To apply the concept of immigration and death processes to the airline booking process, we can think of airline bookings as a population. Each request for a reservation is a potential new immigrant to the population. Importantly, the arrival rate of additional requests for reservations *does not* depend on the number of bookings currently in the population. Because of advance purchase requirements on various fares and the fact that low fare travelers tend to book early with high fare travelers booking late, the arrival rate of a particular fare class may vary during the time period before the flight departs. A go-show can be characterized as a last minute request for a reservation.

On the other hand, cancellations are deaths of existing members of the population. The number of cancellations per period depends on the number of bookings currently in the population. The cancellation rate may vary during the time period before the flight departs. Cancellations tend to be high at three points during the booking process: early in the booking process because of changes in travel plans, at the time of the advance purchase limit because the airline automatically cancels the bookings which are not ticketed in compliance with the restrictions, and on the day of departure because of no-shows. At other times during the booking process, the cancellation rate is usually lower. In this probabilistic context, a no-show is viewed as a last minute cancellation.

If we assume the *Markov property*, that the probability of any particular future behavior of the process, given its current state, is not altered by additional knowledge concerning its past behavior, then we can represent the airline booking process as an *immigration-death process.* The Markov property corresponds to the assumption that future bookings on a particular flight depend only on the current number of bookings on-hand, not on how the bookings on-hand were generated. Most airline demand modeling experts (most notably Littlewood, 1972), believe that the Markov property holds for airline booking data. In the next two subsections, we formulate the airline booking process as an immigration and death process and examine the following cases:

Case 1: Single fare class, infinite capacity

Case 2: Single fare class, finite capacity

4.4.1 Case 1: Single Fare Class, Infinite Capacity

In the single fare class, infinite capacity case, the aircraft is assumed to be "elastic". That is, no booking limits are imposed and all of the requests which arrive for each fare class are accepted. Now, we define the terminology and state the assumptions of this model. Let $B_{cfd}(\tau)$ be the number of bookings in fare class c on flight f departing on date d on-hand at time τ days after the airline has started accepting reservations. In the following discussion, we omit the subscripts c, f, and d for the sake of clarity. We distinguish time t and time τ. When t is used, time is counted back from the day of departure. Time $t = 0$ is the day of departure and time $t = M$ is M days before departure, when the airline starts accepting reservations. In this chapter, time τ is used and measures time forward from the start of the booking process to the day of departure. Time $\tau = 0$ is when the airline starts accepting reservations for this flight and $\tau = M$ is the day of departure, M days after the start of the booking process. For example, if the airline starts accepting reservations 365 days before departure, M = 365.

Requests for reservations are assumed to arrive in a Poisson manner with a time-dependent rate $\lambda(\tau)$, $0 \le \tau \le M$, independent of population size. Thus, in a very small period of time $\Delta\tau$, the probability of an arrival is $\lambda(\tau)\Delta\tau$. Reservations are subject to random cancellation, in the sense that the probability of any individual reservation being cancelled in time $\Delta\tau$ is $\mu(\tau)\Delta\tau$, where $\mu(\tau)$ is the time-dependent cancellation rate. When there are n bookings in the system, the expected number of cancellations occurring in the whole population in the interval $\Delta\tau$ is $n\mu(\tau)\Delta\tau$. Finally, the state of the booking process at the starting time is known. In order to treat the most general case, we assume that there are B(0) bookings on-hand when the booking process starts at time $\tau = 0$.

In order to write the conditional probabilities describing the stochastic process, we assume that the period $\Delta\tau$ is very short, such that at most one reservation, one cancellation, or nothing at all occurs. Then, we have the following conditional probabilities:

$$P[B(\tau+\Delta\tau) = n+1 \mid B(\tau) = n] = \lambda(\tau)\Delta\tau + o_1(\Delta\tau) \tag{4.1a}$$

$$P[B(\tau+\Delta\tau) = n-1 \mid B(\tau) = n] = n\mu(\tau)\Delta\tau + o_2(\Delta\tau) \tag{4.1b}$$

$$P[B(\tau+\Delta\tau) = n \mid B(\tau) = n] = 1 - \lambda(\tau)\Delta\tau - n\mu(\tau)\Delta\tau + o_3(\Delta\tau) \tag{4.1c}$$

$$P[B(\tau+\Delta\tau) = k \mid B(\tau) = n] = o_4(\Delta\tau) \quad \text{for } |k-n| > 1 \tag{4.1d}$$

where $o_i(\Delta\tau)$ for i = 1, 2, 3, and 4 represents higher order terms such that $\lim_{\Delta\tau \to 0} \left[\frac{o_i(\Delta\tau)}{\Delta\tau} \right] = 0$ and $\sum_{i=1}^{4} o_i(\Delta\tau) = 0$.

Let $P_n(\tau) \equiv P[B(\tau) = n]$ for notational convenience. Then, given the equations (4.1) and ignoring the higher order terms $o_i(\Delta\tau)$, we can write:

$$P_n(\tau+\Delta\tau) = P_{n+1}(\tau)[(n+1)\mu(\tau)\Delta\tau] + P_n(\tau)[1 - \lambda(\tau)\Delta\tau - n\mu(\tau)\Delta\tau] + P_{n-1}(\tau)\,[\lambda(\tau)\Delta\tau]\,,\ n \geq 0 \quad (4.2)$$

where we define $P_{-1} = 0$. Equation (4.2) is valid since, if $n > 0$, the state involving precisely n bookings in the interval $[0, \tau + \Delta\tau]$ arises from n -1 bookings in $[0, \tau]$ with 1 request in time $\Delta\tau$, *or* from n +1 bookings in $[0, \tau]$ with 1 cancellation in time $\Delta\tau$, *or* from n bookings in $[0, \tau]$ and nothing occurring in $\Delta\tau$. Rearranging terms in (4.2) and dividing by $\Delta\tau$, we obtain:

$$\frac{P_n(\tau+\Delta\tau) - P_n(\tau)}{\Delta\tau} = -(\lambda(\tau) + n\mu(\tau))P_n(\tau) + (n+1)\mu(\tau)P_{n+1}(\tau) + \lambda(\tau)P_{n-1}(\tau)\,,\ n \geq 0 \qquad (4.3)$$

As $\Delta\tau \to 0$, we obtain the following differential equation which describes the single class, infinite capacity airline booking process:

$$\frac{dP_n(\tau)}{d\tau} = -(\lambda(\tau) + n\mu(\tau))P_n(\tau) + (n+1)\mu(\tau)P_{n+1}(\tau) + \lambda(\tau)P_{n-1}(\tau)\,,\ n \geq 0 \qquad (4.4)$$

with initial condition $P_m(0) = \begin{cases} 1 & \text{if } m = B(0) \\ 0 & \text{otherwise} \end{cases}$ since we start with B(0) bookings in the system at time 0 during the booking process.

We want to solve equation (4.4) to find an expression for $P_n(\tau)$. Bailey (1964) suggests the method of probability generating functions to solve equations similar to (4.4). We define the probability generating function in the usual manner:

$$G(z,\tau) = \sum_{n=0}^{\infty} P_n(\tau)\, z^n \tag{4.5}$$

To apply the generating function method to the airline booking process, we multiply equation (4.4) by z^n and sum from zero to infinity. This gives us:

$$\sum_{n=0}^{\infty} \frac{dP_n(\tau)}{d\tau} z^n = -\sum_{n=0}^{\infty} (\lambda(\tau)+n\mu(\tau))P_n(\tau)z^n + \sum_{n=0}^{\infty} (n+1)\mu(\tau)P_{n+1}(\tau)z^n + \sum_{n=0}^{\infty} \lambda(\tau)P_{n-1}(\tau)z^n \tag{4.6}$$

The second step of the solution method is to rewrite equation (4.6) in terms of $G(z,\tau)$ and its partial derivatives. We determine the partial derivatives of $G(z,\tau)$ with respect to z and τ:

$$\begin{aligned} \frac{\partial G(z,\tau)}{\partial \tau} &= \frac{\partial}{\partial \tau}\Big(\sum_{n=0}^{\infty} P_n(\tau)z^n\Big) \\ &= \sum_{n=0}^{\infty} \frac{\partial P_n(\tau)}{\partial \tau} z^n \quad \text{since } z \text{ is a constant with respect to } \tau \end{aligned} \tag{4.7}$$

$$\frac{\partial G(z,\tau)}{\partial z} = \frac{\partial}{\partial z}\Big(\sum_{n=0}^{\infty} P_n(\tau)z^n\Big)$$

$$= \sum_{n=0}^{\infty} P_n(\tau) \frac{\partial z^n}{\partial z}$$

$$= \sum_{n=0}^{\infty} P_n(\tau)\, n z^{n-1} \tag{4.8}$$

Substituting $G(z,\tau)$ and its partial derivatives (4.5), (4.7), and (4.8) into equation (4.6), we obtain:

$$\frac{\partial G(z,\tau)}{\partial \tau} = -\lambda(\tau)G(z,\tau) - \mu(\tau)\sum_{n=0}^{\infty} nP_n(\tau)z^n + \mu(\tau)\sum_{n=0}^{\infty}(n+1)P_{n+1}(\tau)z^n$$

$$+ \lambda(\tau)\sum_{n=1}^{\infty} P_{n-1}(\tau)z^n$$

$$= -\lambda(\tau)G(z,\tau) - \mu(\tau)z\sum_{n=0}^{\infty} nP_n(\tau)z^{n-1} + \mu(\tau)\sum_{n=0}^{\infty} P_{n+1}(\tau)\frac{\partial z^{n+1}}{\partial z}$$

$$+ z\lambda(\tau)\sum_{n=1}^{\infty} P_{n-1}(\tau)z^{n-1}$$

$$= -\lambda(\tau)G(z,\tau) - \mu(\tau)z\frac{\partial G(z,\tau)}{\partial z} + \mu(\tau)\frac{\partial G(z,\tau)}{\partial z} + \lambda(\tau)zG(z,\tau)$$

using standard change of variable techniques. If we combine terms in the above equation, we have:

$$\frac{\partial G(z,\tau)}{\partial \tau} = \lambda(\tau)(z-1)G(z,\tau) - \mu(\tau)(z-1)\frac{\partial G(z,\tau)}{\partial z} \tag{4.9}$$

with initial condition $G(z,0) = z^{B(0)}$, which arises because we allow B(0) bookings in the system at the start of the process. In order to solve equation (4.9), we need the following theorem (Kells,1975) on the solution of first order partial differential equations.

Theorem 4.1

If $u(z,\tau,G) = C_1$ and $v(z,\tau,G) = C_2$ are two independent integrals of the auxiliary ordinary differential equations

$$\frac{dz}{P(G,z,\tau)} = \frac{d\tau}{Q(G,z,\tau)} = \frac{dG}{R(G,z,\tau)}$$

then

$$C_2 = \psi(C_1)$$

is the unique general solution of

$$Pp + Qq = R.$$

Proof: See Kells (1975) or Ford (1955).

Q. E. D.

The following theorem uses the solution technique given in Theorem 4.1 to solve equation (4.9).

Theorem 4.2

The unique solution to the partial differential equation (4.9) is

$$G(z,\tau) = \left[1 + (z-1)\exp\left(-\int_0^\tau \mu(\alpha)d\alpha\right)\right]^{B(0)} \exp\left[(z-1)\int_0^\tau \lambda(s)\exp\left(-\int_s^\tau \mu(\alpha)d\alpha\right)ds\right]$$

Proof:

In order to find the unique solution of the partial differential equation (4.9), we use the "auxiliary equations" method given in Theorem 4.1. Note that equation (4.9) is in the form $R = Pp + Qq$ required in Theorem 4.1, if we let $P = \mu(\tau)(z-1)$, $Q = 1$, $R = \lambda(\tau)(z-1)G(z,\tau)$, $p = \frac{\partial G(z,\tau)}{\partial z}$, and $q = \frac{\partial G(z,\tau)}{\partial \tau}$. For notational simplicity, we will let $G = G(z,\tau)$. The auxiliary equations of (4.9) are formulated as follows:

$$\frac{dG}{\lambda(\tau)(z-1)G} = \frac{dz}{\mu(\tau)(z-1)} = \frac{d\tau}{1} \tag{4.10}$$

The solution method proceeds by taking two independent integrals of (4.10). The first independent integral is:

$$\frac{dz}{\mu(\tau)(z-1)} = \frac{d\tau}{1}$$

or, equivalently, we can write $\mu(\tau)d\tau = \frac{dz}{(z-1)}$. Integrating both sides of the equation, we obtain:

$$\int_0^\tau \mu(\alpha)d\alpha + C = \ln(z-1)$$

where C is the constant of integration. Then, we exponentiate both sides and get:

$$z-1 = C_1 \exp\left(\int_0^\tau \mu(\alpha)d\alpha\right) \tag{4.11}$$

where $C_1 = \exp(C)$. Equivalently,

$$C_1 = (z-1)\exp\left(-\int_0^\tau \mu(\alpha)d\alpha\right) \tag{4.12}$$

A second independent integral of (4.10) is:

$$\frac{dG}{\lambda(\tau)(z-1)G} = \frac{d\tau}{1}$$

or, equivalently, we get:

$$\lambda(\tau)(z-1)d\tau = \frac{dG}{G}$$

Substituting for (z–1) from equation (4.11), we have:

$$\lambda(\tau)C_1 \exp(\int_0^{\tau}\mu(\alpha)d\alpha)d\tau = \frac{dG}{G}$$

We integrate both sides and obtain:

$$C_1 \int_0^{\tau} \lambda(s)\exp(\int_0^{s}\mu(\alpha)d\alpha)ds + C_2 = \ln(G) \tag{4.13}$$

Then, we substitute (4.12) into (4.13) and the resulting equation is:

$$\ln(G) = \left[(z-1)\ \exp(-\int_0^{\tau}\mu(\alpha)d\alpha)\right]\int_0^{\tau} \lambda(s)\exp(\int_0^{s}\mu(\alpha)d\alpha)ds\ +\ C_2 \tag{4.14}$$

The final step of the solution procedure for (4.9) is to find the most general solution of (4.9). Theorem 4.1 states that the most general solution is of the form:

$$C_2 = \psi(C_1) \tag{4.15}$$

where ψ is determined by the initial conditions. We substitute (4.12) and (4.14) into (4.15):

$$\ln(G) - \left[(z-1)\ \exp(-\int_0^{\tau}\mu(\alpha)d\alpha)\right]\int_0^{\tau} \lambda(s)\exp(\int_0^{s}\mu(\alpha)d\alpha)ds\ = \psi((z-1)\ \exp(-\int_0^{\tau}\mu(\alpha)d\alpha))$$

We exponentiate both sides and obtain:

$$G = \exp\left[\psi\left((z-1)\exp\left(-\int_0^\tau \mu(\alpha)d\alpha\right)\right)\right] * \exp\left[(z-1)\exp\left(-\int_0^\tau \mu(\alpha)d\alpha\right)\int_0^\tau \lambda(s)\exp\left(\int_0^s \mu(\alpha)d\alpha\right)ds\right] \tag{4.16}$$

In equation (4.16), we can simplify the rightmost expression.

$$\begin{aligned}
\exp\left(-\int_0^\tau \mu(\alpha)d\alpha\right)\int_0^\tau \lambda(s)\exp\left(\int_0^s \mu(\alpha)d\alpha\right)ds &= \int_0^\tau \lambda(s)\exp\left(-\int_0^\tau \mu(\alpha)d\alpha\right)\exp\left(\int_0^s \mu(\alpha)d\alpha\right)ds \\
&= \int_0^\tau \lambda(s)\exp\left(-\int_s^\tau \mu(\alpha)d\alpha\right)ds
\end{aligned}$$

since $\exp\left(-\int_0^\tau \mu(\alpha)d\alpha\right)$ is constant with respect to s. Hence, a slightly simplified version of (4.16) is the following:

$$G = \exp\left[\psi\left((z-1)\exp\left(-\int_0^\tau \mu(\alpha)d\alpha\right)\right)\right] * \exp\left[(z-1)\int_0^\tau \lambda(s)\exp\left(-\int_s^\tau \mu(\alpha)d\alpha\right)ds\right] \tag{4.17}$$

To find the particular solution of (4.9), we recall that the initial conditions state that $G(z,0) = z^{B(0)}$. So, setting $\tau = 0$ in (4.17), we have the following relationship:

$$\exp[\psi(z-1)] = G(z,0) = z^{B(0)} \tag{4.18}$$

Let u = z−1. Then, equation (4.18) becomes:

$$\exp[\psi(u)] = (u+1)^{B(0)}$$

or, equivalently, taking the ln of both sides,

$$\psi(u) = B(0)\ln(u+1) \tag{4.19}$$

Equation (4.19) gives us the necessary functional form of ψ. Substituting into (4.17), we obtain the particular solution to (4.9):

$$G(z,\tau) = \exp\left[B(0)^{*}\ln\left(1 + (z-1)\exp(-\int_0^{\tau}\mu(\alpha)d\alpha)\right)\right]^{*}\exp\left[(z-1)\int_0^{\tau}\lambda(s)\exp(-\int_s^{\tau}\mu(\alpha)d\alpha)ds\right]$$

$$= \left[1 + (z-1)\exp(-\int_0^{\tau}\mu(\alpha)d\alpha)\right]^{B(0)}\exp\left[(z-1)\int_0^{\tau}\lambda(s)\exp(-\int_s^{\tau}\mu(\alpha)d\alpha)ds\right] \tag{4.20}$$

Q. E. D.

Equation (4.20) is the probability generating function of the distribution of the number of bookings in the reservations system at time τ after the booking process has started. There are several interesting observations to be made about equation (4.20). First, we note that $G(z,\tau)$ is the product of a Binomial probability generating function and a Poisson probability generating function. The Binomial probability generating function has the form $(1 + (z-1)p)^n$, where p is the probability of "success" and n is the total number of trials. (Medhi, 1982) From (4.20), we have $p = \exp(-\int_0^{\tau}\mu(\alpha)d\alpha)$ and $n = B(0)$. This Binomial probability refers to the number of surviving bookings of the original B(0) in the reservations system, where p refers to the probability of survival and n is the number of bookings initially in the reservations system. The probability that an original booking survives τ time units is $\exp(-\int_0^\tau \mu(\alpha)d\alpha)$.

The Poisson probability generating function has the form $\exp(a(z-1))$, where a is the Poisson "arrival" rate. (Medhi, 1982) In equation (4.20), $a = \int_0^\tau \lambda(s)\exp(-\int_s^\tau \mu(\alpha)d\alpha)ds$. This Poisson probability describes the number of new reservations arriving into the reservations system, where a is the arrival rate. Note that the arrival rate a on the interval $[0,\tau]$ is the product of the request rate, $\lambda(s)$, at any time s and the probability of surviving the remainder of the interval, $\exp(-\int_s^\tau \mu(\alpha)d\alpha)$, integrated over the interval $[0,\tau]$.

When there are initially no bookings in the system, B(0) = 0 and equation (4.20) simplifies considerably. The following corollary states the result.

Corollary 4.1.1

When there are no bookings in the system initially, equation (4.20) simplifies to:

$$G(z,\tau) = \exp\left[(z-1) \int_0^\tau \lambda(s)\exp(-\int_s^\tau \mu(\alpha)d\alpha)ds \right] \tag{4.21}$$

Proof:

Substitute $B(0) = 0$ into (4.20). The leftmost term on the righthand side of (4.20) vanishes and we are left with equation (4.21).

Q. E. D.

Equation (4.21) corresponds to a Poisson distribution with arrival rate

$$\int_0^{\tau} \lambda(s)\exp\left(-\int_s^{\tau}\mu(\alpha)d\alpha\right)ds.$$

Now, we want to calculate the expected number of bookings at any time τ after the beginning of the booking process. The following corollary establishes the desired result.

Corollary 4.1.2

At time τ after the booking process starts, the expected number of bookings is

$$E[B(\tau)|B(0)] = B(0)^{*}\exp\left(-\int_0^{\tau}\mu(\alpha)d\alpha\right) + \int_0^{\tau}\lambda(s)\exp\left(-\int_s^{\tau}\mu(\alpha)d\alpha\right)ds.$$

Proof:

The expected number of bookings at time τ after the start of the booking process is obtained from the probability generating function by taking the derivative of (4.20) with respect to z, evaluated at $z = 1$. We have:

$$\frac{\partial G(z,\tau)}{\partial z} = \exp\left[(z-1)\int_0^\tau \lambda(s)\exp(-\int_s^\tau \mu(\alpha)d\alpha)ds\right] * B(0)\left[1 + (z-1)\exp(-\int_0^\tau \mu(\alpha)d\alpha)\right]^{B(0)-1}$$

$$* \exp(-\int_0^\tau \mu(\alpha)d\alpha)$$

$$+ (1 + (z-1)\exp(-\int_0^\tau \mu(\alpha)d\alpha))^{B(0)} * \exp\left[(z-1)\int_0^\tau \lambda(s)\exp(-\int_s^\tau \mu(\alpha)d\alpha)ds\right]$$

$$* \int_0^\tau \lambda(s)\exp(-\int_s^\tau \mu(\alpha)d\alpha)ds \qquad (4.22)$$

Evaluating equation (4.22) at $z = 1$, we obtain the expected number of bookings at time τ given the initial number of bookings at time 0:

$$E[B(\tau)|B(0)] = (\exp(0)^* B(0)^* \exp(-\int_0^\tau \mu(\alpha)d\alpha)) + (1^* \exp(0)^* \int_0^\tau \lambda(s)\exp(-\int_s^\tau \mu(\alpha)d\alpha)ds)$$

$$= B(0)^* \exp(-\int_0^\tau \mu(\alpha)d\alpha) + \int_0^\tau \lambda(s)\exp(-\int_s^\tau \mu(\alpha)d\alpha)ds \qquad (4.23a)$$

Q. E. D.

If we let $z(\tau) = \exp(-\int_0^\tau \mu(\alpha)d\alpha)$ and $a(\tau) = \int_0^\tau \lambda(s)\exp(-\int_s^\tau \mu(\alpha)d\alpha)ds$, then equation (4.23a) becomes

$$E[B(\tau)|B(0)] = B(0)^* z(\tau) + a(\tau) \qquad (4.23b)$$

Note that $z(\tau)$ is the probability that an initial reservation remains in the system and $a(\tau)$ is the expected number of new bookings in the system between the start of the booking process and time τ. Equation (4.23a) allows an intuitive explanation of the expected number of bookings in the system at time τ. The expected number of bookings at time τ equals the expected number of the initial bookings remaining in the system plus the expected number of new bookings obtained between time 0 and time τ.

It is important to observe that, because of the Markov assumption, (4.23a) holds on any interval during the booking process. Suppose that an airline is currently at time τ days after the start of the booking process of a particular flight and wants to forecast the expected number of bookings in class c at time M days after the start of the booking process (the day of departure). Thus, the airline wants to forecast $E[B(M)|B(\tau)]$. From (4.23a), we get:

$$E[B(M)|B(\tau)] = B(\tau)^{*}\exp\left(-\int_{\tau}^{M}\mu(\alpha)d\alpha\right) + \int_{\tau}^{M}\lambda(s)\exp\left(-\int_{s}^{M}\mu(\alpha)d\alpha\right)ds \tag{4.24}$$

Another observation about the immigration and death model of the airline booking process is that, when the request rate $\lambda(\tau)$ and cancellation rate $\mu(\tau)$ are homogeneous over time, the probability generating function (4.20) simplifies considerably. Corolllary 4.1.3 states the result.

Corollary 4.1.3

When $\lambda(\tau) = \lambda$ and $\mu(\tau) = \mu$, the probability generating function (4.20) becomes

$$G(z,\tau) = [1 + (z-1)\exp(-\mu\tau)]^{B(0)} \exp\left[(z-1)\frac{\lambda}{\mu}[1 - \exp(-\mu\tau)]\right].$$

Proof:

We substitute $\lambda(\tau) = \lambda$ and $\mu(\tau) = \mu$ into (4.20) and we get:

$$\begin{aligned} G(z,\tau) &= \left[1 + (z-1)\exp(-\int_0^\tau \mu d\alpha)\right]^{B(0)} \exp\left[(z-1)\int_0^\tau \lambda\exp(-\int_s^\tau \mu d\alpha)ds\right] \\ &= [1 + (z-1)\exp(-\mu\tau)]^{B(0)} \exp\left[(z-1)\,\lambda\int_0^\tau \exp(-\mu(\tau-s))ds\right] \\ &= [1 + (z-1)\exp(-\mu\tau)]^{B(0)} \exp\left[(z-1)\frac{\lambda}{\mu}[1 - \exp(-\mu\tau)]\right] \end{aligned} \tag{4.25}$$

Q. E. D.

For the homogeneous request and cancellation rate case, Goel and Richter-Dyn (1974) state a closed form solution for $P_n(\tau)$:

$$P_n(\tau) = \exp\left\{-\frac{\lambda}{\mu}[1 - \exp(-\mu\tau)]\right\} \sum_{k=0}^{\min(j,n)} \frac{\binom{j}{k}\exp(-\mu\tau k)(1-\exp(-\mu\tau))^{j+n-2k}}{(n-k)!}\left(\frac{\lambda}{\mu}\right)^{n-k} \tag{4.26}$$

where $j = B(0)$.

Recall that the probability generating function (4.20) is the product of a Poisson probability generating function and a Binomial probability generating function. We can apply the Poisson approximation to the Binomial distribution to equation (4.20). The result is established in the following corollary.

<u>Corollary 4.1.4</u>

If the Poisson approximation to the Binomial distribution is applied to equation (4.20), we obtain the following approximation to $G(z,\tau)$:

$$H(z,\tau) = \exp\left[(z-1)\left(B(0)\exp(-\int_0^\tau \mu(\alpha)d\alpha) + \int_0^\tau \lambda(s)\exp(-\int_s^\tau \mu(\alpha)d\alpha)ds \right)\right].$$

<u>Proof</u>:

The Poisson approximation to the Binomial distribution is stated as follows (Larsen and Marx, 1986): as $n \to \infty$, a Binomial distribution with parameters n and p approaches a Poisson distribution with parameter $a = np$. In terms of probability generating functions, we have the following relationship (as B(0) becomes large):

$$\left[1 + (z-1)\exp(-\int_0^\tau \mu(\alpha)d\alpha)\right]^{B(0)} \approx \exp\left[(z-1)B(0)\exp(-\int_0^\tau \mu(\alpha)d\alpha)ds\right]$$

Substituting this expression into (4.20), we find $H(z,\tau)$ as an approximation to $G(z,\tau)$:

$$H(z,\tau) = \exp\left[(z-1)B(0)\exp(-\int_0^\tau \mu(\alpha)d\alpha)\right]\exp\left[(z-1)\int_0^\tau \lambda(s)\exp(-\int_s^\tau \mu(\alpha)d\alpha)ds\right]$$

$$= \exp\left[(z-1)\left(B(0)\exp(-\int_0^\tau \mu(\alpha)d\alpha) + \int_0^\tau \lambda(s)\exp(-\int_s^\tau \mu(\alpha)d\alpha)ds\right)\right] \qquad (4.27)$$

Q. E. D.

After applying the Poisson approximation, we observe that (4.27) is the probability generating function for the Poisson distribution with arrival rate $m(\tau)$, where

$$m(\tau) = B(0)\exp(-\int_0^\tau \mu(\alpha)d\alpha) + \int_0^\tau \lambda(s)\exp(-\int_s^\tau \mu(\alpha)d\alpha)ds$$

We have approximated the immigration and death process by an immigration (Poisson) process with a modified arrival rate. In effect, we increase the arrival rate to account for the survivors of the initial population. In airline booking terminology, the request and cancellation process can be approximated by a request process with a modified request rate. When the request rate $\lambda(\tau)$ and the cancellation rate $\mu(\tau)$ are homogeneous over time ($\lambda(\tau) = \lambda$ and $\mu(\tau) = \mu$), we apply the Poisson approximation to equation (4.25):

$$H(z,\tau) = \exp[(z-1)B(0)\exp(-\mu\tau)]\exp\left[(z-1)\frac{\lambda}{\mu}[1-\exp(-\mu\tau)]\right]$$

$$= \exp\left\{(z-1)\left[\frac{\lambda}{\mu} + \left(B(0) - \frac{\lambda}{\mu}\right)\exp(-\mu\tau)\right]\right\} \qquad (4.28)$$

which, as above, is a Poisson process with a modified request rate.

4.4.2 Case 2: Single Fare Class, Finite Capacity

In case 2, we assume a single fare class with finite capacity. The airline places a booking limit on each fare class, reflecting the finite capacity of the aircraft and the airline's very short run revenue maximization goal. At any time τ days after the booking process begins, the booking limit of fare class c, $CAP(\tau)$, is known and fixed. The booking limit constrains the state space of the stochastic process to the set of non-negative integers less than or equal to $CAP(\tau)$.

As in case 1, $B(\tau)$ is the number of bookings in fare class c on flight f departing on date d on-hand at time τ days after the airline has started accepting reservations. Requests for reservations are assumed to arrive in a Poisson manner with time-dependent rate $\lambda(\tau)$, $0 \leq \tau \leq M$, independent of population size. However, because of the booking limit, the request rate $\lambda(\tau)$ falls to 0 when the number of bookings is greater than $CAP(\tau)$. Cancellations occur in a random manner, where $\mu(\tau)$ is the time-dependent cancellation rate.

We require that initial bookings be less than the booking limit at any time τ, $B(0) \leq \min_{\tau}(CAP(\tau))$, to insure proper operation of the stochastic process. It is important to understand that, occasionally, total bookings may exceed the booking limit. For example, if an airline seat inventory control analyst decides that "too many" spaces have already been sold, the analyst may set the booking limit below current total bookings so that no further bookings are allowed. In this section, we do not treat this special case. However, it will be discussed as an extension in Chapter 8.

Now, we are able to write the conditional probabilities for the finite capacity case. We assume that the period $\Delta\tau$ is very short, so that at most one request, one cancellation, or nothing at all occurs. The conditional probabilities are:

$$P[B(\tau+\Delta\tau) = n+1 \mid B(\tau) = n] = \begin{cases} \lambda(\tau)\Delta\tau + o_1(\Delta\tau), \; n = 0,1,\ldots,CAP(\tau)-1 \\ 0 \text{ otherwise} \end{cases} \tag{4.29a}$$

$$P[B(\tau+\Delta\tau) = n-1 \mid B(\tau) = n] = \begin{cases} n\mu(\tau)\Delta\tau + o_2(\Delta\tau), \; n = 0,1,\ldots, CAP(\tau) \\ 0 \text{ otherwise} \end{cases} \tag{4.29b}$$

$$P[B(\tau+\Delta\tau) = n \mid B(\tau) = n] = \begin{cases} 1 - \lambda(\tau)\Delta\tau - n\mu(\tau)\Delta\tau + o_3(\Delta\tau), \; n = 0,1,\ldots,CAP(\tau)-1 \\ 1 - n\mu(\tau)\Delta\tau + o_3(\Delta\tau), \; n = CAP(\tau) \\ 0 \text{ otherwise} \end{cases} \tag{4.29c}$$

$$P[B(\tau+\Delta\tau) = k \mid B(\tau) = n] = \begin{cases} o_4(\Delta\tau) \text{ for } \mid k - n \mid > 1, \; n = 0,1,\ldots, CAP(\tau) \\ 0 \text{ otherwise} \end{cases} \tag{4.29d}$$

where $o_i(\Delta\tau)$ represents higher order terms such that $\lim_{\Delta\tau\to 0}\left[\frac{o_i(\Delta\tau)}{\Delta\tau}\right] = 0$ and $\sum_{i=1}^{4} o_i(\Delta\tau) = 0$.

If we let $P_n(\tau) = P[B(\tau) = n]$, then the differential equations describing the airline booking process for $n = 0,1,\ldots,CAP(\tau)-1$ are the same as in case 1. For $n = 0,1,\ldots,CAP(\tau)-1$, equations (4.29) reduce to equations (4.1) and, hence, the differential equation (4.4) still holds:

$$\frac{dP_n(\tau)}{d\tau} = -(\lambda(\tau) + n\mu(\tau))P_n(\tau) + (n+1)\mu(\tau)P_{n+1}(\tau) + \lambda(\tau)P_{n-1}(\tau), \; 0 \le n \le CAP(\tau)-1 \tag{4.30}$$

For $n = CAP(\tau)$, we can write the following equation (ignoring higher order terms):

$$P_n(\tau+\Delta\tau) = P_{n-1}(\tau)\lambda(\tau)\Delta\tau + P_n(\tau)(1-n\mu(\tau)\Delta\tau) \qquad (4.31)$$

Equation (4.31) holds true since, for $n = CAP(\tau)$, the state involving exactly n bookings in the interval $[0, \tau+\Delta\tau]$ is obtained from $n-1$ bookings in $[0,\tau]$ and one request in time $\Delta\tau$ *or* from n bookings in $[0,\tau]$ and nothing happening in time $\Delta\tau$. The booking limit does not allow the booking process to reach state CAP+1.

We rearrange terms in (4.31) and divide by $\Delta\tau$:

$$\frac{P_n(\tau+\Delta\tau) - P_n(\tau)}{\Delta\tau} = \lambda(\tau)P_{n-1}(\tau) - P_n(\tau)n\mu(\tau)$$

As $\Delta\tau \rightarrow 0$, we have the following differential equation for $n = CAP(\tau)$:

$$\frac{dP_n(\tau)}{d\tau} = \lambda(\tau)P_{n-1}(\tau) - P_n(\tau)n\mu(\tau) \qquad (4.32)$$

The initial condition is $P_m(0) = \begin{cases} 1 & \text{if } m = B(0) \\ 0 & \text{otherwise} \end{cases}$ since we start with B(0) bookings in the system at time 0 of the booking process.

We want to solve the system of differential equations formed by (4.30) and (4.32). However, this is a difficult system of equations to solve, because of the time-dependent rates $\lambda(\tau)$ and $\mu(\tau)$ and the time-dependent booking limit $CAP(\tau)$. Larson and Odoni (1981) suggest numerical procedures to solve this type of system of differential equations. Unfortunately, numerical procedures require the unknown parameters $\lambda(\tau)$ and $\mu(\tau)$ to be specified.

Let us assume that the request rate, the cancellation rate, and the booking limit are homogeneous over time, where $\lambda(\tau) = \lambda$, $\mu(\tau) = \mu$, and $CAP(\tau) = CAP$ for all τ, equations (4.30) and (4.32) become:

$$\frac{dP_n(\tau)}{d\tau} = -(\lambda + n\mu)P_n(\tau) + (n+1)\mu P_{n+1}(\tau) + \lambda P_{n-1}(\tau),\ 0 \leq n \leq CAP-1 \tag{4.33}$$

$$\frac{dP_n(\tau)}{d\tau} = P_{n-1}(\tau)\lambda(\tau) - P_n(\tau)n\mu(\tau),\quad n = CAP \tag{4.34}$$

with initial condition $P_m(0) = \begin{cases} 1 & \text{if } m = B(0) \\ 0 & \text{otherwise} \end{cases}$, $B(0) \leq CAP$. Theorem 4.3 gives the solution to equations (4.33) and (4.34).

Theorem 4.3

The solution to the system of differential equations (4.33) and (4.34) is, for n = 0,1,...,CAP,

$$P_n(\tau) = \frac{\dfrac{\rho^n e^{-\rho}}{n!}}{\displaystyle\sum_{l=0}^{CAP} \frac{\rho^l e^{-\rho}}{l!}} + \frac{CAP!}{n!}\rho^{CAP-B(0)} \sum_r \frac{D_{B(0)}(r)D_n(r)}{rD_{CAP}(r)D'_{CAP}(r+1)} \exp(r\mu\tau)$$

(4.35)

where r sums over the c roots of $D_c(s+1) = 0$, $\rho = \frac{\lambda}{\mu}$, $D_i(r) = \sum_{w=0}^{i} \binom{i}{w} \rho^{i-w}\, r(r+1)\ldots(r+w-1)$,

and $D'_n(r+1) = \sum_{w=1}^{n} \frac{(n)_w}{w} D_{n-w}(r+1)$,.

Proof: Riordan (1962).

It is interesting to note that the leftmost term on the righthand side of (4.35) forms a truncated Poisson distribution with mean $\rho = \frac{\lambda}{\mu}$ and truncation from above at the booking limit. Note that the truncated Poisson distribution is independent of the initial condition B(0). This implies that the rightmost term on the righthand side of (4.35) is the transient part of $P_n(\tau)$. Equation (4.35) is a very complex probability statement.

A second approach to the finite capacity case is to consider a constrained version of the Poisson approximation to the infinite capacity case. Recall equation (4.27), which is the probability generating function of a Poisson process with a modified request rate:

$$H(z,\tau) = \exp\left[(z-1)\left(B(0)\exp(-\int_0^\tau \mu(\alpha)d\alpha) + \int_0^\tau \lambda(s)\exp(-\int_s^\tau \mu(\alpha)d\alpha)ds\right)\right] \tag{4.36}$$

Let $m(\tau) = B(0)\exp(-\int_0^\tau \mu(\alpha)d\alpha) + \int_0^\tau \lambda(s)\exp(-\int_s^\tau \mu(\alpha)d\alpha)ds$. Then, inverting the probability generating function (4.36), we obtain the following state probabilities for the infinite capacity case:

$$P_n(\tau) = \exp(-m(\tau)) \frac{m(\tau)^n}{n!}, \; n = 0,1,2,... \tag{4.37}$$

where $m(\tau)$ is defined as above. Now, suppose there is a booking limit, $CAP(\tau)$, placed on the Poisson process given in equation (4.37). The booking limit constrains the Poisson process (4.37) at $CAP(\tau)$. Thus, for $n < CAP(\tau)$, the state probabilities remain as in equation (4.37). However, at $CAP(\tau)$, we observe a spike which represents all the probability of reaching states greater than or equal to the booking limit. In other words, we have the following probability distribution for the number of bookings at time τ days after the start of the booking process:

$$P_n(\tau) = \exp(-m(\tau)) \frac{m(\tau)^n}{n!}, \; n = 0,1,2,...,CAP(\tau)-1 \tag{4.38a}$$

$$P_n(\tau) = \sum_{j=n}^{\infty} \exp(-m(\tau)) \frac{m(\tau)^j}{j!}, \; n = CAP(\tau) \tag{4.38b}$$

where $P_n(\tau) = 0$ for $n > CAP(\tau)$ and $m(\tau)$ is defined as above. Equations (4.38) form a *censored* Poisson distribution, censored from above at the booking limit, $CAP(\tau)$.

4.5 Applicability of Stochastic Models to Forecasting

This section discusses the applicability of the stochastic models developed in the previous section to forecasting the true, underlying demand for airline reservations. We make the distinction between what an airline observes in its reservations system and what the airline wants to know in order to maximize revenue. First, the airline observes a booking process for each class which is constrained at a known booking limit. Thus, the appropriate model for the observed process is the single fare class, *finite* capacity model. Second, for the revenue maximization procedure, an airline wants to know the true, unconstrained demand. In effect, the airline wants an estimate of the demand that would occur if it had an "elastic" aircraft and no booking limits existed. Therefore, the appropriate model is the single fare class, *infinite* capacity model.

The preceding discussion suggests an intuitive algorithm for finding the true, underlying demand for bookings. The algorithm starts by realizing that historical data on previous departures of the same flight number is taken from a finite capacity booking process. As a result, the algorithm statistically estimates the parameters $\lambda(\tau)$ and $\mu(\tau)$ from one of the finite capacity models using available historical data. However, the airline wants to determine the unconstrained demand. Thus, the second step of the algorithm is to obtain a forecast of expected bookings for a future flight by substituting the estimated parameters into the infinite capacity model. In particular, we substitute the estimated parameters $\lambda(\tau)$ and $\mu(\tau)$ from the finite capacity model into the expression for expected bookings to come from the infinite capacity model.

4.6 Estimation and Forecasting in the Stochastic Models

Estimation of the parameters of the stochastic models introduced in section 4.4 can be difficult because of the complex nature of the probability statements for $P_n(\tau)$, the probability of n bookings in the reservations system at time t after the start of the booking process, in the finite capacity case. After the estimation step, forecasting the true, unconstrained demand is a straightforward substitution of the estimated parameters into an expected value statement derived from the infinite capacity case. In this section, we propose a maximum likelihood estimation procedure for estimating the parameters in the finite capacity case. Then, for forecasting purposes, we derive the proper expected value statement from the infinite capacity case.

In the finite capacity case, we found that equations (4.30) and (4.32) are mathematically intractable. There is no known solution for the system of differential equations which describe the single fare class, finite capacity case with non-homogeneous request and cancellation rates. (Larson and Odoni, 1981) Then, we assumed constant request and cancellation rates. Although this assumption allows us to find a closed form solution for the probability of n bookings in the reservations system at time τ after the start of the booking process, the probabilities defined by (4.35) are mathematically complex. It is, of course, theoretically possible to apply maximum likelihood estimation and estimate the parameters λ and μ. However, we would prefer a mathematically more tractable probability statement.

Recall the constrained version of the Poisson approximation to the infinite capacity case given by equations (4.38). Equations (4.38) formed a censored Poisson distribution, censored from above at the booking limit. The censored Poisson distribution is mathematically tractable and maximum likelihood estimation can be applied in a straightforward manner. The primary drawback to equations (4.38) is the need to specify $\lambda(\tau)$ and $\mu(\tau)$. A simple and reasonable specification is to assume that the arrival and cancellation rates are constant on time intervals during the booking process. Several major U.S. airlines identify such time intervals, varying by fare class and destination. An additional feature is that airlines generally do not modify the booking limits during these intervals. Therefore, we will also assume constant booking limits on each time interval.

Let the booking process start at time $\tau = \tau_0$ and the flight depart at time $\tau = \tau_M$ after the start of the booking process. Suppose that we split the time interval $[\tau_0, \tau_M]$ into L subintervals $[\tau_0, \tau_1]$, $(\tau_1, \tau_2]$, $(\tau_2, \tau_3]$, ... , $(\tau_{L-1}, \tau_L]$, where $\tau_L = \tau_M$. We assume that, on any interval $(\tau_l, \tau_{l+1}]$, $\lambda(\tau) = \lambda_l$ and $\mu(\tau) = \mu_l$, and $CAP(\tau) = CAP_l$. Then, according to the Markov property, the reservations and cancellations in any subinterval $[\tau_l, \tau_{l+1}]$ depend only on the state of the system at time τ_l and the parameters λ_l and μ_l.

Suppose we have observations at each of the L+1 time points during the booking process: $B(0) = b_0$, $B(\tau_1) = b_1$, $B(\tau_2) = b_2$, ... , $B(\tau_{L-1}) = b_{L-1}$, $B(\tau_L) = b_L$. The following theorem shows that the joint probability $P[B(\tau_L) = b_L, B(\tau_{L-1}) = b_{L-1}, \ldots, B(\tau_2) = b_2, B(\tau_1) = b_1, B(0) = b_0]$ equals the product of the probabilities on each subinterval, given the number of bookings at the end of the previous interval.

<u>Theorem 4.4</u>

Suppose that we divide the time interval $[\tau_0, \tau_M]$ into L subintervals $[\tau_0, \tau_1]$, $[\tau_1, \tau_2]$, $[\tau_2, \tau_3]$, ... , $[\tau_{L-1}, \tau_L]$, where $\tau_L = \tau_M$. We assume that, on any interval $[\tau_l, \tau_{l+1}]$, $\lambda(\tau) = \lambda_l$ and $\mu(\tau) = \mu_l$, and $CAP(\tau) = CAP_l$. If $P[B(\tau_{l+1}) = b_{l+1} \mid B(\tau_l) = b_l]$ for l = 0, 1, ... , L−1 is given by the censored Poisson distribution in equation (4.38), then

$$P[B(\tau_L) = b_L, B(\tau_{L-1}) = b_{L-1}, \ldots, B(\tau_2) = b_2, B(\tau_1) = b_1, B(0) = b_0] =$$

$$\prod_{j=0}^{L} P[B(\tau_{L-j}) = b_{L-j} \mid B(\tau_{L-j-1}) = b_{L-j-1}]$$

Proof:

By elementary probability laws for conditional probabilities, we have

$$
\begin{aligned}
&P[B(\tau_L) = b_L, B(\tau_{L-1}) = b_{L-1}, \dots, B(\tau_2) = b_2, B(\tau_1) = b_1, B(0) = b_0] = \\
&P[B(\tau_L) = b_L | B(\tau_{L-1}) = b_{L-1}, \dots, B(\tau_2) = b_2, B(\tau_1) = b_1, B(0) = b_0] * \\
&P[B(\tau_{L-1}) = b_{L-1} \mid B(\tau_{L-2}) = b_{L-2}, \dots, B(\tau_1) = b_1, B(0) = b_0] * \\
&* \dots * \\
&P[B(\tau_2) = b_2 \mid B(\tau_1) = b_1, B(0) = b_0] * \\
&P[B(\tau_1) = b_1 | B(0) = b_0) * P[B(0) = b_0]
\end{aligned}
$$

We can simplify this equation by appealing to the Markov property of the censored Poisson distribution. That is, for any arbitrary interval $[\tau_l, \tau_{l+1}]$, the probability of b_{l+1} bookings at time t_{l+1} depends only on the number of bookings at time t_l. Hence, the above probability statement simplifies to:

$$
\begin{aligned}
&P[B(\tau_L) = b_L, B(\tau_{L-1}) = b_{L-1}, \dots, B(\tau_2) = b_2, B(\tau_1) = b_1, B(0) = b_0] \\
&\quad = P[B(\tau_L) = b_L | B(\tau_{L-1}) = b_{L-1}] * P[B(\tau_{L-1}) = b_{L-1} \mid B(\tau_{L-2}) = b_{L-2}] * \dots * \\
&\quad\quad P[B(\tau_2) = b_2 \mid B(\tau_1) = b_1] * P[B(\tau_1) = b_1 | B(0) = b_0] * P[B(0) = b_0] \\
&\quad = \prod_{j=0}^{L} P[B(\tau_{L-j}) = b_{L-j} \mid B(\tau_{L-j-1}) = b_{L-j-1}] \qquad (4.39)
\end{aligned}
$$

where equation (4.39) gives us the desired result.

Q. E. D.

Recall that $P[B(\tau_{l+1}) = b_{l+1} \mid B(\tau_l) = b_l]$ for $l = 0, 1, \ldots, L-1$ in equation (4.39) is given by the censored Poisson distribution in equation (4.38). We note that, on the subinterval $[\tau_l, \tau_{l+1}]$ for any l, the conditional probability statement for $B(\tau_{l+1})$ includes $B(\tau_l)$. Since $B(\tau_l)$ is known, the likelihood function corresponding to equation (4.39) is separable by period. Thus, maximum likelihood estimates of λ_l and μ_l for all periods l are found as follows:

1. Identify subintervals for each fare class on which the booking limit as well as the request and cancellation rates are approximately constant.

2. Estimate the parameters λ_l and μ_l on each of the L subintervals via maximum likelihood estimation starting with $[\tau_0, \tau_1]$ and ending with $[\tau_{L-1}, \tau_L]$.

Finally, we need to develop an expected value statement for the infinite capacity case, in order to forecast the true, unconstrained demand for bookings. Suppose that an airline is at time τ_i after the start of the booking process and that the airline wants to determine the expected number of bookings on the day of departure (time τ_M). In statistical terms, the airline wants to find $E[B(\tau_M)|B(\tau_i)]$ and the following theorem establishes the result.

Theorem 4.5

If the request rate and cancellation rate is constant on time intervals, then the expected bookings on the day of departure τ_M given that the booking process is currently at time τ_i is

$$E[B(\tau_M)|B(\tau_i)] = B(\tau_i)^*\exp\left[\sum_{l=i}^{L-1} \mu_l(\tau_l-\tau_{l+1})\right]$$

$$+ \sum_{j=i}^{L-2} \frac{\lambda_j}{\mu_j} \exp\left(\sum_{k=j+1}^{L-1} \mu_k(\tau_k-\tau_{k+1})\right)\left(1-\exp\left(\mu_j(\tau_j-\tau_{j+1})\right)\right)$$

$$+ \frac{\lambda_{L-1}}{\mu_{L-1}} \left(1-\exp\left(\mu_{L-1}(\tau_{L-1}-\tau_L)\right)\right)$$

where we let $\tau_L = \tau_M$.

Proof:

In the infinite capacity case, equation (4.24) gives the appropriate expected value:

$$E[B(\tau_M)|B(\tau_i)] = B(\tau_i)^*\exp(-\int_{\tau_i}^{\tau_L}\mu(\alpha)d\alpha) + \int_{\tau_i}^{\tau_L}\lambda(s)\exp(-\int_{s}^{\tau_L}\mu(\alpha)d\alpha)ds \tag{4.40}$$

But $\lambda(\tau)$ and $\mu(\tau)$ are constant on each subinterval. Thus, the integrals in (4.40) simplify and we have

$$\exp\left(-\int_{\tau_i}^{\tau_L}\mu(\alpha)d\alpha\right) = \exp\left(-\int_{\tau_i}^{\tau_{i+1}}\mu_i d\alpha - \int_{\tau_{i+1}}^{\tau_{i+2}}\mu_{i+1}d\alpha - \ldots - \int_{\tau_{L-1}}^{\tau_L}\mu_{L-1}d\alpha\right)$$

$$= \exp(\mu_i(\tau_i - \tau_{i+1}) + \mu_{i+1}(\tau_{i+1} - \tau_{i+2}) + \ldots + \mu_{L-1}(\tau_{L-1} - \tau_L))$$

$$= \exp\left[\sum_{l=i}^{L-1}\mu_l(\tau_l - \tau_{l+1})\right] \tag{4.41}$$

and

$$\int_{\tau_i}^{\tau_L}\lambda(s)\exp\left(-\int_s^{\tau_L}\mu(\alpha)d\alpha\right)ds = \int_{\tau_i}^{\tau_{i+1}}\lambda_i\exp\left(-\int_s^{\tau_L}\mu(\alpha)d\alpha\right)ds + \int_{\tau_{i+1}}^{\tau_{i+2}}\lambda_{i+1}\exp\left(-\int_s^{\tau_L}\mu(\alpha)d\alpha\right)ds + \ldots + \int_{\tau_{L-1}}^{\tau_L}\lambda_{L-1}\exp\left(-\int_s^{\tau_L}\mu(\alpha)d\alpha\right)ds \tag{4.42}$$

Equation (4.42) can be simplified by substituting (4.41) into the inner integral and evaluating the integral. For example, the first term on the righthand side of (4.42) becomes:

$$\int_{\tau_i}^{\tau_{i+1}}\lambda_i\exp\left(-\int_s^{\tau_L}\mu(\alpha)d\alpha\right)ds = \int_{\tau_i}^{\tau_{i+1}}\lambda_i\exp\left(-\int_s^{\tau_{i+1}}\mu_i d\alpha - \int_{\tau_{i+1}}^{\tau_{i+2}}\mu_{i+1}d\alpha - \ldots - \int_{\tau_{L-1}}^{\tau_L}\mu_{L-1}d\alpha\right)ds$$

$$= \int_{\tau_i}^{\tau_{i+1}}\lambda_i\exp(\mu_i(s - \tau_{i+1}) + \mu_{i+1}(\tau_{i+1} - \tau_{i+2}) + \ldots + \mu_{L-1}(\tau_{L-1} - \tau_L))ds$$

$$= \frac{\lambda_j}{\mu_j} \exp\left(\sum_{k=j+1}^{L-1} \mu_k(\tau_k - \tau_{k+1}) - \sum_{k=j}^{L-1} \mu_k(\tau_k - \tau_{k+1}) \right)$$

$$= \frac{\lambda_j}{\mu_j} \exp\left(\sum_{k=j+1}^{L-1} \mu_k(\tau_k - \tau_{k+1}) \right) \left(1 - \exp\left(\mu_j(\tau_j - \tau_{j+1})\right)\right) \tag{4.43}$$

If we substitute (4.41), (4.42), and (4.43) into (4.40), we have the desired expected value expression:

$$E[B(\tau_M)|B(\tau_i)] = B(\tau_i)^* \exp\left[\sum_{l=i}^{L-1} \mu_l(\tau_l - \tau_{l+1}) \right]$$

$$+ \sum_{j=i}^{L-2} \frac{\lambda_j}{\mu_j} \exp\left(\sum_{k=j+1}^{L-1} \mu_k(\tau_k - \tau_{k+1}) \right) \left(1 - \exp\left(\mu_j(\tau_j - \tau_{j+1})\right)\right)$$

$$+ \frac{\lambda_{L-1}}{\mu_{L-1}} \left(1 - \exp\left(\mu_{L-1}(\tau_{L-1} - \tau_L)\right)\right) \tag{4.44}$$

Q. E. D.

The leftmost term on the righthand side of equation (4.44) is the expected number of the initial bookings at time τ_i remaining in the system at time τ_M. The two remaining terms on the righthand side of equation (4.44) represent the expected number of bookings obtained in each subinterval which still remain at time τ_M, summed over all of the time intervals. To forecast the expected number of bookings from time τ_i after the start of the booking process to the day of departure (time τ_M), we substitute the estimated parameters λ_l and μ_l for all l into equation (4.44).

4.7 The "Ideal" Stochastic Model

This section describes the "ideal" stochastic model, attempting to extend the stochastic model presented in this chapter to include more of the actual airline booking process as presented in Chapter 2. The extended model captures many of the characteristics of the actual booking process as an airline "sees" it. Theoretically, most of these characteristics can be taken into account in a stochastic model. However, the resulting system of differential equations describing the system would most likely be mathematically intractable. As a result, the discussion in this section is qualitative in nature. The goal is to present a more general stochastic model. We will examine the request (arrival) process, the cancellation (departure) process, and the state space (number of bookings).

The arrival process consists of requests for reservations in a particular fare class on a certain flight departing on a specific date. For any particular fare class, the request rate is *non-homogeneous* over the duration of the booking process. For example, fare classes containing fares used primarily by leisure travelers exhibit high request rates early in the booking process and lower request rates late in the booking process. A go-show can be regarded as a last minute request. A second feature is that the request process is an *immigration* process. in general, requests for a particular flight enter the booking process from outside sources, usually from the population surrounding the market. Bookings currently in the system do not generate additional requests. As a result, the request rate is generally independent of the state of the system.

Another aspect of the request process is how many requests arrive in a single instant. Earlier in the chapter, we assumed that only one request arrives in a single instant of time. Tour operators and travel agents often form tour groups to vacation destinations and, hence, request a large number of seats at the same time. In stochastic modeling terminology, group requests are considered *bulk arrivals* into the booking process. It is important to note that group requests at a single point in time can be approximated by a series of single requests over a short period of time.

Since most airlines assign several different fare classes to the Economy cabin, the request process is truly a multivariate process. That is, requests for different classes are competing for the same inventory of seats. In a multivariate setting, it is possible to examine the dependencies between demand in different fare classes. Many of these dependencies arise because of the limited number of seats allotted to each fare class. If the requested fare class is closed, then the request is *denied*.

When a request is denied, at least one of the three following actions occurs: waitlisting, retrial, or a decision not to travel. Waitlisting causes a *queue* to form and allows the stochastic process to be modeled using queueing theory. If a traveler decides to request a different fare class on the same flight (vertical movement), a different flight on the same airline (horizontal movement), or requests the desired fare class on the same flight at a later time during the booking process, the traveler makes a *retrial* to the booking process. That is, the request rate at any time may reflect previously denied requests (as opposed to new demand). Finally, a decision not to travel on the airline represent a *loss* to the booking process.

The departure process consists of cancellations of existing reservations. Cancellations may be made by the passenger due to a change in travel plans or by the airline when a traveler fails to meet established deadlines for ticketing. No-shows can be regarded as last minute cancellations. The cancellation rate may be *non-homogeneous* over the duration of the booking process. For example, a certain fare may allow changes up to 3 days before departure. It would be plausible to expect a higher cancellation rate 4 or 5 days before departure than earlier in the booking process.

The cancellation rate may also be *age-dependent*. In effect, the cancellation rate may depend on how long the reservation has been in the system. Some airline researchers advance the "bathtub" shape of the cancellation rate for an individual reservation. At first, the cancellation rate is high because of uncertainty of travel plans. Then, when the reservation is of medium age, the cancellation rate is low. If there is a time limit for ticketing placed on the reservation, the cancellation rate will be high at that point. Finally, as the day of departure draws near, the cancellation rate increases due to last minute changes in travel plans or no-shows.

Earlier in the chapter, we assumed that only one cancellation is possible at a single point in time. As with requests, we may observe *bulk departures* from the system. We can approximate bulk departures by treating them as a series of cancellations over a short period of time. Finally, the cancellation process is a *death* process. That is, the number of cancellations which occur during a given time period depends on the number of bookings in the reservations system.

The final topic to examine is the state space. The state space consists of the spaces available in a certain fare class on a particular flight. Although we can calculate an effective booking limit for each class and apply the single fare class, finite capacity models discussed earlier in the chapter, the booking process is a multivariate process which consists of a *shared* inventory of spaces. There are restrictions on which fare classes can occupy which spaces. If there are four fare classes (where fare class 1 has the highest fares and class 4 has the lowest), the state space is a *nested* inventory of spaces. Some spaces can only be occupied by class 1 reservations, other spaces with class 1 and 2 reservations, another set of spaces with class 1, 2, and 3 reservations, and a final set is available to all 4 classes. A multivariate model would allow us to capture the interactions between requests and cancellations in the various fare classes and the shared, nested inventory of spaces.

Overall, this section has tried to capture qualitatively some of the more subtle aspects of the booking process. As we saw earlier in this chapter, relatively simple assumptions lead to mathematically complex probability expressions. However, some of the more complicated assumptions can be tested in simulations and incorporated into models as needed. It should be noted that, as the complexity of the stochastic model increases, the quantity of data required for parameter estimation increases dramatically.

4.8 Conclusions and Implications

This chapter has developed a probabilistic model for the analysis of the booking process. After introducing the assumption that demand for each fare class is distinct and stationary, we model the booking process as an immigration and death process with non-homogeneous request and cancellation rates. The resulting state probabilities are mathematically complex. Therefore, we make some simplifying assumptions and demonstrate that the booking process can be modeled as a censored Poisson process. We can estimate the request and cancellation rates via maximum likelihood estimation on the single fare class, finite capacity model. A forecast of true, unconstrained demand for bookings on the day of departure is found by substituting the estimated rates into an expected value expression derived from the single fare class, infinite capacity model.

For the airlines, these probabilistic models can provide insight into the underlying request and cancellation rates and lead to more informed seat inventory control decisions. The principal drawback is the necessity of using maximum likelihood estimation, which can be time consuming and costly. Major U.S. airlines are, from the practical standpoint, attempting to forecast for several thousand flights per day over a period of up to one year before departure. The prospect of performing a censored Poisson regression for each forecast on every flight may be unrealistic. However, the applicability of the probabilistic models hinges on the ability to identify the flights which require the most attention, paticularly the high demand flights. The need for accurate forecasts on high demand flights is particularly critical -- as one or two spaces on each high demand flight sold at a higher fare can provide significant additional revenue on an annual basis.

Furthermore, the probabilistic models can provide insight into the booking process through simulation methodology. The inherently dynamic nature of the booking process makes it difficult, if not impossible, for even the most experienced seat inventory control analyst to fully understand all of the implications of a particular inventory decision. Simulation based on the ideal stochastic model could provide invaluable "what if?" capabilities for an airline. For example, the effect of changing booking limits at particular points during the booking process could be studied. Also, tour groups cause much frustration for seat inventory control analysts because of the propensity of group sizes to shrink or entire groups to cancel. Simulation of the booking process might provide additional information about the need for tour groups and the effect of tour group cancellation rates on the booking process. In conclusion, probabilistic modeling has definite potential for significant contributions in the yield management area at major airlines.

Chapter 5 A Statistical Framework for Analysis of the Booking Process

5.1 Introduction

In the previous chapter, we developed a rigorous probabilistic model of the airline booking process. The probabilistic model emphasized the dynamic and stochastic nature of the booking process, where requests and cancellations occur at non-homogeneous rates throughout the weeks and months before a flight departs. This chapter focuses on a comprehensive statistical framework of the airline booking process. In essence, we treat the booking process from a data analysis perspective. The emphasis in this chapter is on capturing the patterns in the booking data.

This chapter begins with a description of the basic definitions and terminology of the booking process from the statistical viewpoint. Next, we investigate the available airline data and distinguish between estimation and forecasting of the airline booking process. Three types of statistical models naturally arise from the data: advance bookings, historical bookings, and combined models. The advance bookings model is derived from bookings already made on a particular flight; the historical bookings model arises from bookings on previous departures of the same flight number; and the combined models are a combination of the two approaches. We show that the combined model provides an intuitive view of the booking process.

The second part of this chapter addresses the issue of booking limits and demand distributions. First, the issue of booking limits is considered. Booking limits are crucially important since the airline wants to estimate the true, underlying consumer demand. The total number of bookings observed in each class reflects the true underlying consumer demand constrained by a set of nested booking limits. The presence of booking limits leads to a truncated–censored demand distribution and a set of truncated–censored regression equations. Finally, several possible demand distributions, such as the Normal, Log-normal, Poisson and Gamma, are introduced.

5.2 Terminology of the Booking Process

As described in Chapter 2, the airline booking process is quite complex. In the weeks before a flight departs, many reservations are made in each fare class. Interspersed with additional reservations are cancellations of previously made reservations. Furthermore, go-shows and no-shows add additional complications into the booking process. Finally, external factors such as fare levels for each class, flight frequency, season of the year, aircraft type, and frequent traveler award plans may have a significant impact on the booking process for a particular flight.

The booking process measures the net bookings made during a given time period for a specific class, flight, and date. Net bookings equal the additional reservations made minus cancellations of previously made reservations. That is, for any given time period,

NET BOOKINGS = ADDITIONAL RESERVATIONS - CANCELLATIONS

For example, the time period of interest can be one day. Therefore, net bookings generated at time t refer to net bookings generated between day t+1 before departure and day t before departure. Formally, we define $b_{cfd}(t)$ as the net bookings generated at time t (t days before departure) in fare class c on flight f departing on date d. In this chapter, we note that the time t is counted back from the day of departure. Thus, time t = 0 is the day of departure and time t = M is M days before departure, when the airline starts accepting reservations.

A cumulative measure of the net reservations made up to the current time can now be defined. Let total bookings, $B_{cfd}(t)$, at time t days before departure in fare class c on flight f departing on date d be the cumulative number of net bookings from the start of the booking process (say M days before the departure date) to the current time t days before departure. For the mathematical development, we assume that a single day is small relative to the entire booking horizon (M days). So, we can integrate over the net bookings generated at each time s to obtain total bookings. Thus,

$$B_{cfd}(t) = \int_{s=M}^{t} b_{cfd}(s)\, ds \qquad (5.1)$$

The booking curve is the curve traced out by $B_{cfd}(t)$ as we vary t from t = M (the start date of the booking process) to t = 0 (the departure date of the flight). For example, $B_{cfd}(M+1) = 0$ since reservations are not accepted more than M days before departure. On the other hand, $B_{cfd}(0)$ is the total number of bookings in fare class c on-hand at departure time. Recall that Figure 2.3 shows a sample booking curve.

We define bookings to come, $BTC_{cfd}(t)$, at time t days before departure in fare class c on flight f departing on date d as the net reservations made from t days before departure to the departure date of the flight. Hence, we have

$$BTC_{cfd}(t) = \int_{s=t}^{0} b_{cfd}(s)\, ds = B_{cfd}(0) - B_{cfd}(t) \tag{5.2}$$

Bookings to come is a very important concept because, at any time t days before departure, an airline knows how many bookings are on-hand from its computer reservations systems. Therefore, at that point in time, an airline is interested in forecasting bookings to come for each class on every flight.

An important concept is the distinction between the number of bookings at departure time, $B_{cfd}(0)$, and the actual number of passengers boarded on the aircraft. The actual number of passengers boarded is determined not only by the number of bookings, but also by the number of travelers with reservations who do not show up or have missed connecting flights, the number of travelers who show up without reservations, the number who are waitlisted for the flight, and the number of travelers with valid tickets but have no recorded reservations. Also, we note that an aircraft has a maximum capacity, no matter how many potential travelers show up.

To formalize these concepts, we define the following terms:

1. Waitlisted travelers ($WL_{cfd}(t)$) are the potential travelers on the waiting list at time t days before departure for fare class c on flight f departing on date d.

2. No-shows (NS_{cfd}) are the passengers holding reservations who do not show up for fare class c on flight f departing on date d. No-shows are caused by passengers holding reservations who do not present themselves available at the airport at departure time. An important type of no-show is a misconnect. A misconnect is a passenger with reservations who is not able to show up for fare class c on flight f departing on date d because of a missed connection. Misconnects are caused by the airline's failure to deliver passengers to a hub airport in time to make their scheduled connecting flight.

3. Go-shows (GS_{cfd}) are the travelers, not previously waitlisted, who show up without reservations and without a valid ticket for fare class c on flight f departing on date d.

4. No-recs (NR_{cfd}) are the travelers holding valid tickets for fare class c on flight f departing on date d, but have no reservation recorded in the reservations system.

5. Capacity ($CAP_{cfd}(t)$) is the maximum authorized (or physical) capacity of fare class c at time t days before departure on flight f departing on date d. Note that, due to overbooking,the maximum authorized capacity at time t days before departure may be greater than the physical capacity of the aircraft. On the other hand, aircraft range and weight limitations may force the maximum authorized capacity to be less than the physical capacity of the aircraft. At departure time, the maximum authorized capacity is no more than the physical capacity.

6. Passengers boarded (P_{cfd}) are the passengers who board the aircraft in fare class c on flight f departing on date d.

Mathematically, we have the following relationship:

$$P_{cfd} = MIN[B_{cfd}(0) + WL_{cfd}(0) + GS_{cfd} + NR_{cfd} - NS_{cfd}, CAP_{cfd}(0)] \qquad (5.3)$$

5.3 Available Booking Data

We examine the available booking data from a typical airline data base. Table 5.1 shows the data available at date d on a particular flight f for fare class c. In this table, we suppress the subscripts f and c for clarity. Examining Table 5.1, two types of data are available for statistical modeling:

1. Advance bookings - bookings already made on the particular flight for which a bookings forecast is required.

2. Historical bookings - bookings at various points in time prior to departure on previous departures of the same flight number, perhaps categorized by day of week or season.

$B_d(t)$ is the total number of bookings at time t (days before departure) departing on date d. P_d represents the total passengers boarded on the flight departing on date d. Advance bookings are represented as *columns* of Table 5.1. For example, the column for departure date d–2 of Table 5.1 traces out the complete booking curve of the flight which departed on date d-2. The column corresponding to departure date d+2 traces out the partial booking curve of the flight which will depart on date d+2. Historical bookings are the *rows* of Table 5.1. For instance, the second row of Table 5.1 contains bookings on the day of departure (time $t = 0$) for each departure date. The question marks on Table 5.1 represent values that are not yet known on date d.

Departure Dates →

···	Date d–2	Date d–1	**Date d**	Date d+1	Date d+2	···	Date d+M	
···	P_{d-2}	P_{d-1}	P_d	?	?	?	?	Boarded
···	$B_{d-2}(0)$	$B_{d-1}(0)$	$B_d(0)$	?	?	?	?	Day 0
···	$B_{d-2}(1)$	$B_{d-1}(1)$	$B_d(1)$	$B_{d+1}(1)$	?	?	?	
···	$B_{d-2}(2)$	$B_{d-1}(2)$	$B_d(2)$	$B_{d+1}(2)$	$B_{d+2}(2)$	?	?	**Time**
···	$B_{d-2}(3)$	$B_{d-1}(3)$	$B_d(3)$	$B_{d+1}(3)$	$B_{d+2}(3)$	···	?	**Before**
···	$B_{d-2}(4)$	$B_{d-1}(4)$	$B_d(4)$	$B_{d+1}(4)$	$B_{d+2}(4)$	···	?	**Departure**
.	.	.	.	.	.	.	.	
.	.	.	.	.	.	.	.	↓
.	.	.	.	.	.	.	.	
···	$B_{d-2}(M-1)$	$B_{d-1}(M-1)$	$B_d(M-1)$	$B_{d+1}(M-1)$	$B_{d+2}(M-1)$	···	?	Day M–1
···	$B_{d-2}(M)$	$B_{d-1}(M)$	$B_d(M)$	$B_{d+1}(M)$	$B_{d+2}(M)$	···	$B_{d+M}(M)$	Day M

Table 5.1 Available Booking Data in Class c on Flight f on Date d

5.4 Distinction Between Estimation and Forecasting

The airline's goal is to forecast total bookings on the day of departure for each fare class c, flight f, and date d. This forecast is required in the days, weeks, or months before the departure of the flight. In order to produce an accurate forecast, there is a two step process: estimation and forecasting. First, based on past and current data on-hand, we estimate a statistical model. The estimation phase fits a model to the known past observations. Once a model is fit to the data, we enter the forecasting phase. In this phase, we predict future (unknown) values given the current data on-hand. Producing an accurate forecast (or prediction) requires use of the estimated model and intelligent extrapolation. Thus, as the three main statistical models are introduced in the next section, we will explain their use in both estimation and forecasting.

5.5 Three Main Statistical Models

In this section, we will introduce three basic models for estimating and forecasting the booking process. First, we present the fundamental relationship between total bookings on the day of departure and total bookings on-hand at time t days before departure. Then, we describe the three models for the prediction of total bookings on the day of departure: advance bookings, historical bookings, and combined models. We define $B_d(0)$ as the total bookings on the day of departure for a particular flight f and fare class c on a given date d and $B_d(t)$ as the bookings at time t (t days before departure) for a flight f departing on date d. We omit the subscripts c and f for simplicity.

5.5.1 Fundamental Relationship

The fundamental relationship relates current on-hand bookings at time t days before departure to total bookings on the day of departure. Suppose an airline desires to predict toal bookings on the day of departure in fare class c on flight f departing on date d. Then, at any arbitrary time t days before departure, a predictor of total bookings on the day of departure is the sum of bookings on-hand at time t days before departure and the bookings to come between time t days before departure and the day of departure (time 0). In mathematical terms, we have:

$$\begin{aligned} B_d(0) &= B_d(t) + BTC_d(t) \\ &= B_d(t) + \int_{s=t}^{0} b_d(\tau)\, ds \quad , \quad t = M, M-1, \ldots, 0 \qquad (5.4) \end{aligned}$$

However, at time t days before departure, the bookings generating function $b_d(s)$ is unknown for s = t to 0. Suppose we estimate $b_d(t)$ with a known generating function $b_d(t,\alpha)$, where α is a vector of parameters to be estimated plus an error term $\varepsilon(t,d)$, which depends on the time t days before departure and the departure date d. Therefore,

$$b_d(t) = b_d(t,\alpha) + \varepsilon(t,d) \qquad (5.5)$$

Substituting (5.5) into (5.4), we obtain an estimate of the bookings to come on flight f:

$$\begin{aligned} B_d(0) &= B_d(t) + \int_{s=t}^{0} (b_d(s,\alpha) + \varepsilon(s,d))ds \\ &= B_d(t) + B_d(0,\alpha) - B_d(t,\alpha) + \int_{s=t}^{0} \varepsilon(s,d)ds \\ &= B_d(t) + [B_d(0,\alpha) - B_d(t,\alpha)] + \varepsilon_{td} \qquad (5.6) \end{aligned}$$

where $B_d(0,\alpha) - B_d(t,\alpha) = \int_{s=t}^{0} b_d(s,\alpha)\, ds$ and $\varepsilon_{td} = \int_{s=t}^{0} \varepsilon(s,d)\, ds$. The term $[B_d(0,\alpha) - B_d(t,\alpha)]$ is the estimate of bookings to come between the time t days before departure and the flight time (day 0), α is a vector of parameters to be estimated, and ε_{td} is an error term which depends on t, the number of days before departure, and the departure date d. Equation (5.6) defines the *fundamental relationship* between total bookings on the day of departure and current on-hand bookings.

If we obtain a statistical estimate α' of the parameters α, then the fundamental relationship (5.6) becomes

$$\hat{B}_d(0) = B_d(t) + B_d(0,\alpha') - B_d(t,\alpha') \tag{5.7}$$

where $\hat{B}_d(0)$ is a forecast of total bookings on the day of departure for class c on flight f departing on date d. Thus, to forecast total bookings on the day of departure (day 0), we apply equation (5.7) directly. The next three sections introduce three types of statistical models which specify $B_d(0,\alpha)$ and $B_d(t,\alpha)$.

5.5.2 The Advance Bookings Model

In the advance bookings model, we consider the bookings already made on the particular flight for which a forecast of bookings is required. The columns of Table 5.1 are the data of interest and represent the development of the booking curve for each flight. For departed flights, the booking curve is already complete whereas, for flights scheduled in the future, the booking curve is only partially complete. Thus, we use past booking curve data in order to predict the incomplete portions of future booking curves. Additional advance booking data is derived from exogenous factors which affect the development of the booking curve. These factors may include fares, market share, seasonality indices, and the authorized capacity of the aircraft. We identify two types of advance bookings models: the synthetic booking curve model and the time series of advance bookings model.

The Synthetic Booking Curve Model

The first type of advance bookings model is the *synthetic booking curve* model, which attempts to describe the shape of the booking curve. The booking curve is, thus, synthesized from an approximation to its shape and other exogenous factors. If we assume an approximation to the shape of the booking curve is given by $g(\mathbf{X}_{td},t,\beta)$, then the synthetic booking curve model can be expressed as follows:

$$B_d(t) = g(\mathbf{X}_{td},t,\beta) + \eta_{td} \ , d = 1, \ldots , D; t = 0, \ldots , M \qquad (5.8)$$

where D is the amount of historical data available in the database, $\mathbf{X}_{td}$ is a vector of exogenous factors which may be a function of the time t days before departure and the departure date d, β is the vector of parameters associated with the approximation to the shape of the booking curve, and η_{td} is a random error term, which depends on the time t days before departure and the departure date d.

The function $g(\mathbf{X}_{td},t,\beta)$ may include a nonlinear function of t, such as $\beta_1(\ln t) + \beta_2$(market share), or a piecewise linear approximation to the shape of the booking curve. The censored Poisson model developed in Chapter 4 can be viewed as a synthetic booking curve model where the dependent variable is Poisson distributed, the function $g(\mathbf{X}_{td},t,\beta) = m(\tau)$ given in equation (4.38) and the parameters to be estimated are $\lambda(\tau)$ and $\mu(\tau)$. Equation (5.8) states that total bookings at time t days before departure depend on the shape of the booking curve, which is a function of exogenous factors and the time t before departure.

To produce a forecast of total bookings on the day of departure, we first use a sample of advance booking data from previous departures of the same flight number to estimate the parameters β using equation (5.8). Let the estimated parameters be denoted as β' and we obtain:

$$\hat{B}_d(t,\beta') = g(\mathbf{X}_{td},t,\beta') \tag{5.8a}$$

Then, we substitute equation (5.8a) into the fundamental relationship (5.7) and we have:

$$\hat{B}_d(0) = B_d(t) + g(\mathbf{X}_{0d},0,\beta') - g(\mathbf{X}_{td},t,\beta') \tag{5.8b}$$

Finally, we apply equation (5.8b) to produce a forecast of total bookings on the day of departure using the values of the exogenous factors at time t days before departure on the particular flight of interest. Note that equation (5.8b) also requires a forecast of the exogenous variables at departure time ($\mathbf{X}_{0d}$).

<u>The Time Series of Advance Bookings Model</u>

The second type of advance bookings model is the *time series of advance bookings* model. This model expresses total bookings at time t days before departure as a time series of total bookings at earlier points (time t+1, t+2, ... , M days before departure) in the booking history of all flights with the same flight number and exogenous factors. Hence,

$$B_d(t) = \sum_{l=1}^{M-t} \gamma_{lt}\, B_d(t+l) + W_{td}g + \nu_{td}\,, \quad d = 1, \ldots, D;\ t = 0, \ldots, M \tag{5.9}$$

where γ_t is a vector of parameters associated with the time series of advance bookings which depend on the time t days before departure, ν_{td} is the random error term, which may depend on the departure date d and the time t before departure, $\mathbf{W}_{td}$ is a vector of exogenous factors which may be a function of the departure date d and the time t days before departure, $\mathbf{g}$ is the vector of parameters associated with the exogenous factors, D is the amount of historical data available in the database, and M is the earliest time before departure at which bookings are accepted. Note that (5.9) describes a multivariate combined time series/regression model.

To produce a forecast of total bookings on departure day, we first use a sample of advance booking data from previous departures of the same flight number to estimate the parameters **g** and γ_t for all t using equation (5.9). Let the estimated parameters be denoted as **g'** and γ'_t and equation (5.9) becomes

$$\hat{B}_d(t,\gamma'_t,\mathbf{g}') = \sum_{l=1}^{M-t} \gamma'_{lt}\, B_d(t+l) \; + \mathbf{W}_{td}\mathbf{g}' \,, \quad d = 1, \dots, D; t = 0, \dots, M \tag{5.9a}$$

Then, we substitute equation (5.9a) into the fundamental relationship (5.7) and we obtain

$$\hat{B}_d(0) = B_d(t) + \sum_{l=1}^{M} \gamma'_{l0}\, B_d(l) \; + \mathbf{W}_{0d}\mathbf{g}' - \Big(\sum_{l=1}^{M-t} \gamma'_{lt}\, B_d(t+l) \; + \mathbf{W}_{td}\mathbf{g}' \Big) \tag{5.9b}$$

In order to apply equation (5.9b) to produce a forecast of total bookings on a particular flight, it is necessary to forecast some missing values. Namely, at time t days before departure on a future flight, the following data are not yet available: $B_d(1)$, $B_d(2)$, ... , $B_d(t-1)$. We forecast these missing values by applying (5.9a) in a sequential manner. The sequential forecasting procedure starts by forecasting $B_d(t-1)$using (5.9a) for t = t -1. Using the forecast of $B_d(t-1)$, we apply equation (5.9a) for t = t - 2 and, thereby, generate a forecast of $B_d(t-2)$. We proceed similarly until we create a forecast for $B_d(1)$. Finally, to produce a forecast of total bookings on a particular flight, we apply equation (5.9b) using both the forecasted intermediate values and the actual advance booking data at time t days before departure for a given future flight. Note that equation (5.9b) also requires a forecast of the exogenous variables, $\mathbf{W}_{0d}$ and the sequential forecasting procedure requires forecasts of $\mathbf{W}_{1d}$, $\mathbf{W}_{2d}$, ... , and $\mathbf{W}_{t-1,d}$.

5.5.3 The Historical Bookings Model

In the historical bookings model, the focus is on the total bookings on previous departures of the same flight. The rows of Table 5.1 are the data of interest and represent historical trends in the data. This data reveals the cyclical trends and seasonal variations. Again, we want to predict bookings on flight date (time 0) for a particular flight on date d, $B_d(0)$. B_{d-D} represents the earliest data available and the value of D depends on the size of the airline's database. An autoregressive time-series (AR) model for bookings at time t days before departure follows immediately from the rows of table 5.1:

$$B_d(t) = \sum_{j=1}^{D} \varphi_{jt} B_{d-j}(t) + u_{dt} \quad , t = 0, \dots , M \tag{5.10}$$

where M is the earliest time before departure at which bookings are accepted, $B_{d-j}(t)$ is total bookings at time t days before departure for the flight departing on date d-j, u_{dt} is an error term which may depend on the departure date d and the time t days before departure, and φ_t is a vector of parameters to be estimated. Equation (5.10) defines the historical bookings model which predicts the total bookings on the day of departure in class c for a given flight f on date d.

The residuals from equation (5.10) should be examined for any apparent patterns. If any non-random patterns are found, a moving average (MA) component is added to the model:

$$u_{dt} = \sum_{i=1}^{N} a_{it} u_{d-i,t} + v_{dt} \tag{5.11}$$

where N < D, N is the number of lagged error terms, and a_t is a vector of parameters to be estimated. The v_{dt} terms should behave like "white noise". Substituting equation (5.11) into (5.10), an autoregressive moving average (ARMA) time series model for historical bookings is:

$$B_d(t) = \sum_{j=1}^{D} \varphi_{jt} B_{d-j}(t) + \sum_{i=1}^{N} a_{it} u_{d-i,t} + v_{dt} \quad , \ t = 0, \ldots , M \tag{5.12}$$

Equation (5.12) is defined as the *historical bookings* model. Note that (5.12) describes a multivariate time series model.

To produce a forecast of total bookings on departure day at time t days before departure, we first estimate (5.12). For example, if we estimate (5.12), it becomes the following:

$$\hat{B}_d(t,\varphi'_t,a'_t) = \sum_{j=1}^{D} \varphi'_{jt} B_{d-j}(t) + \sum_{i=1}^{N} a'_{it} u_{d-i,t} \tag{5.12a}$$

where the estimated parameters are denoted by φ'_t and a'_t. Substituting (5.12a) into the fundamental relationship (5.7), we have

$$\hat{B}_d(0) = B_d(t) + \sum_{j=1}^{D} \varphi'_{j0} B_{d-j}(0) + \sum_{i=1}^{N} a'_{i0} u_{d-i,0} - \left(\sum_{j=1}^{D-d} \varphi'_{jt} B_{d-j}(t) + \sum_{i=1}^{N} a'_{it} u_{d-i,t} \right) \tag{5.12b}$$

In order to apply equation (5.12b) to produce a forecast of total bookings on a given flight, $\hat{B}_d(0)$, it is necessary to forecast some missing values of $B_{d-j}(0)$. In particular, at time t days before departure on a future flight, the following data are not yet available: $B_{d-j}(0)$ for j = 1, ... , t - 1. We forecast these missing values by applying equation (5.12a) in a sequential manner. The sequential forecasting procedure starts by forecasting $B_{d-t+1}(0)$ using (5.12a) for d = d - t +1. Then, using the forecast of $B_{d-t+1}(0)$, we apply equation (5.12a) for d = d - t + 2 to generate a forecast of $B_{d-t+2}(0)$. We continue in a similar manner until we produce a forecast of $B_{d-1}(0)$.

Finally, we substitute the forecasted missing values along with the historical data on-hand into equation (5.12b) to produce the desired forecast of total bookings, $\hat{B}_d(0)$.

5.5.4 Combined Models

While the advance bookings models presented in Section 5.5.2 use the booking curve and exogenous factors to predict total bookings, the historical bookings model described in Section 5.5.3 uses historical patterns of the same flight numbers to predict total bookings. By combining the two types of models, we propose a combined model which takes both the booking curve and the historical booking phenomena into account. The goal of a combined model is increased explanatory power when compared to the advance bookings or the historical bookings models alone. We propose two methods for combining the models: the weighted average model and the full information model.

<u>The Weighted Average Model</u>

The *weighted average* model is simply the weighted average of the two models -- the advance bookings and the historical bookings models. Thus, using the time series of advance bookings model (5.8), the following combined model is formed:

$$B_d(t) = \theta_1\left(\sum_{l=1}^{M-t} \gamma_{lt} B_d(t+l) + W_{td} g + v_{td} \right) + \theta_2\left(\sum_{j=1}^{D} \varphi_{jt} B_{d-j}(t) + \sum_{i=1}^{N} a_{it} u_{d-i,t} + v_{dt} \right),$$

$$t = 0,\ldots,M;\ d = 1,\ldots,D \quad (5.13)$$

We note that each model is multiplied by the appropriate θ. Hence, we denote the new parameters with a tilde (~). For example, $\tilde{\gamma} = \theta_1 * \gamma$ and $\tilde{\varphi} = \theta_2 * \varphi$. So, equation (5.13) becomes the following combined model:

$$B_d(t) = \sum_{l=1}^{M-t} \tilde{\gamma}_{lt} B_d(t+l) + W_{td}\tilde{g} + \tilde{v}_{td} + \sum_{j=1}^{D} \tilde{\varphi}_{jt} B_{d-j}(t) + \sum_{i=1}^{N} \tilde{a}_{it} u_{d-i,t} + \tilde{v}_{dt}, \quad t = 0,\dots,M;\ d = 1,\dots,D \tag{5.14}$$

To produce a forecast of total bookings on the day of departure, we first must estimate the combined model (5.14) on historical data. We denote the estimated parameters as $\tilde{\gamma}'$, $\tilde{g}'$, $\tilde{\varphi}'_t$, and $\tilde{a}'_t$ and equation (5.14) becomes

$$\hat{B}_d(t, \tilde{\gamma}'_t, \tilde{g}', \tilde{\varphi}'_t, \tilde{a}'_t) = \sum_{l=1}^{M-t} \tilde{\gamma}_{lt} B_d(t+l) + W_{td}\tilde{g} + + \sum_{j=1}^{D} \tilde{\varphi}_{jt} B_{d-j}(t) + \sum_{i=1}^{N} \tilde{a}_{it} u_{d-i,t}, \quad t = 0,\dots,M;\ d = 1,\dots,D \tag{5.14a}$$

Then, we substitute (5.14a) into the fundamental relationship (5.7) and we have

$$\begin{aligned} \hat{B}_d(0) = B_d(t) + & \sum_{l=1}^{M} \tilde{\gamma}_{l0} B_d(l) + W_{0d}\tilde{g} + + \sum_{j=1}^{D} \tilde{\varphi}_{j0} B_{d-j}(0) + \sum_{i=1}^{N} \tilde{a}_{i0} u_{d-i,0} \\ & -\left(\sum_{l=1}^{M-t} \tilde{\gamma}_{lt} B_d(t+l) + W_{td}\tilde{g} + + \sum_{j=1}^{D} \tilde{\varphi}_{jt} B_{d-j}(t) + \sum_{i=1}^{N} \tilde{a}_{it} u_{d-i,t} \right) \end{aligned} \tag{5.14b}$$

In order to apply (5.14b) to produce a forecast of total bookings on a given flight, we must forecast some missing intermediate values. Specifically, at time t days before departure on a future flight, the following required data are not available: $B_{d-j}(s)$, for $j = 0, 1, \dots, t-1$ and $s = 0, 1, \dots, t-j-1$. We forecast these missing values by applying equation (5.14a) in a sequential manner. First, we set $j = t-1$. Second, the sequential forecasting procedure forecasts $B_{d-t+1}(s)$

for $s = t - j - 1, t - j - 2, \ldots, 0$ using (5.14a) written for the corresponding d and s. Then, using these forecasts as data, we decrement the index j by 1 and generate sequential forecasts for $B_{d-t+2}(s)$ for $s = t - j - 1, t - j - 2, \ldots, 0$ using (5.14a) written for the corresponding d and s. We continue in the same manner until $j = 0$. The final value forecast from the sequential forecasting procedure is $B_d(0)$, the desired forecast of total bookings on the day of departure for the particular flight of interest. Note that the sequential forecasting procedure requires a forecast of the exogenous variables $W_{0d}, \ldots, W_{t-1,d}$.

On the other hand, we can use the synthetic booking curve model as the advance bookings model. Then, multiplying the advance bookings model by θ_1 and the historical bookings model by θ_2, the combined model becomes

$$B_d(t) = g(X_{td}, t, \tilde{\beta}) + \tilde{\eta}_{td} + \sum_{j=1}^{D} \tilde{\varphi}_{jt} B_{d-j}(t) + \sum_{i=1}^{N} \tilde{a}_{it} u_{d-i,t} + \tilde{v}_{dt},$$

$$t = 0,\ldots,M;\ d = 1,\ldots,D \qquad (5.15)$$

where, as before, $\tilde{\beta} = \theta_1 * \beta$, $\tilde{\varphi} = \theta_2 * \varphi$, and so forth. Note that equation (5.15) assumes that g is linear with repect to β . To produce a forecast of total bookings on the day of departure, we first must estimate the combined model (5.15) using historical data. Let the estimated parameters be denoted as $\tilde{\beta}'$, $\tilde{\varphi}'_t$, and $\tilde{a}'_t$. Then, we substitute (5.15) into the fundamental relationship (5.7) and we obtain

$$\hat{B}_d(0) = B_d(t) + g(X_{0d}, 0, \tilde{\beta}) + \sum_{j=1}^{D} \tilde{\varphi}_{j0} B_{d-j}(0) + \sum_{i=1}^{N} \tilde{a}_{i0} u_{d-i,0}$$
$$- \left(g(X_{td}, t, \tilde{\beta}) + \sum_{j=1}^{D} \tilde{\varphi}_{jt} B_{d-j}(t) + \sum_{i=1}^{N} \tilde{a}_{it} u_{d-i,t}\right) \qquad (5.15a)$$

As before, we must sequentially forecast the required missing intermediate values. Namely, at time t days before departure on a future flight, the following required data are not yet available: $B_{d-j}(s)$, for j = 1, 2, ... , t -1 and s = 0, 1, ... , t - j -1. We forecast these missing values by applying equation (5.15a) in a sequential manner. First, we set j = t -1. Second, the sequential forecasting procedure produces a forecast of $B_{d-t+1}(s)$ for s = t - j -1, t - j -2, ... , 0 using (5.15a) written for the corresponding d and s. Then, using these forecasts as data, we decrease the index j by 1 and generate sequential forecasts for $B_{d-t+2}(s)$ for s = t - j -1, t - j -2, ... , 0 using (5.14a) written for the corresponding d and s. We repeat these steps until j = 1. Finally, to produce a forecast of total bookings on a particular flight, we use the forecasted intermediate booking data along with the historical booking data available at time t days before departure on the future flight of interest. Note that the forecasting procedure requires a forecast of the exogenous variables $X_{0d}, \ldots, X_{t-1,d}$.

The Full Information Model

The second type of combined model is the *full information* model. The full information model is developed by viewing the booking process as a time series of historical bookings. Then, each element of the time series is viewed as the result of a booking curve. This method allows a "natural" interpretation of the booking process from the data in Table 5.1. The full information model uses the rows of Table 5.1 to form a time series. Then, each element of the time series is described as a function of the "booking curve" elements in the column directly beneath it.

In order to develop this model mathematically, we express total bookings at flight time as a time series of historical bookings

$$B_d(0) = \sum_{j=0}^{D} \varphi_{j0} B_{d-j}(0) + n_{d0} \tag{5.16}$$

Suppose we are at time t days before departure on flight f departing on date d. Then, we know the total bookings at flight time, $B_{d-j}(0)$, for $j = t,\ldots,D$. For $j = 0,\ldots,t-1$, the flights have not yet departed and we have only partial booking information available. In particular, at time t days before departure, the most recent data available is $B_{d-j}(t-j)$ for $j = 0,\ldots,t-1$. Thus, for $j = 0,\ldots,t-1$, we estimate the total bookings at flight time, $B_{d-j}(0)$, by using the fundamental relationship (5.7). Writing equation (5.7) for $d = d-j$ and $t = t-j$, we obtain:

$$B_{d-j}(0) = B_{d-j}(t-j) + B_{d-j}(0,\alpha) - B_{d-j}(t-j,\alpha) + \varepsilon_{t-j,d} \tag{5.16a}$$

Now, we substitute equation (5.16a) into (5.16) for $j = 0, \ldots, t-1$. The resulting equation is

$$B_d(0) = \sum_{j=t}^{D} \varphi_{j0} B_{d-j}(0) + \Big(\sum_{j=0}^{t-1} \varphi_{j0}(B_{d-j}(t-j)+B_{d-j}(0,\alpha)-B_{d-j}(t-j,\alpha) + \varepsilon_{t-j,d})\Big) + n_{d0} \tag{5.17}$$

Equation (5.17) defines the full information combined model. The final step is to use the a statistical model to describe the bookings to come $B_{d-j}(0,\alpha) - B_{d-j}(t-j,\alpha)$ term. Two potential models are the synthetic booking curve model and a historical moving average of bookings to come.

As an example, we substitute the synthetic booking curve model given by equation (5.8) into (5.17) and obtain the following equation:

$$B_d(0) = \sum_{j=t}^{D} \varphi_{j0} B_{d-j}(0) + \Big(\sum_{j=0}^{t-1} \varphi_{j,t-j}(B_{d-j}(t-j) + g(\mathbf{X}_{0,d-j},0,\beta) - g(\mathbf{X}_{t-j,d-j},t-j,\beta) + \eta_{t-j,0,d-j})\Big) + n_{d0}$$

(5.18)

where $\eta_{t-j,0,d-j} = \eta_{0,d-j} - \eta_{t-j,d-j}$. If we let $n^{\circ} = (\varphi * \eta) + n.$, then (5.18) becomes the following:

$$B_d(0) = \sum_{j=t}^{D} \varphi_{j0} B_{d-j}(0) + \Big(\sum_{j=0}^{t-1} \varphi_{j,t-j}(B_{d-j}(t-j) + g(X_{0,d-j},0,\beta) - g(X_{t-j,d-j},t-j,\beta))\Big) + n^{\circ}{}_{d0}$$

(5.19)

In order to estimate the parameters of equation (5.19), we perform maximum likelihood estimation using advance and historical booking data from previous departures of the same flight number. After obtaining parameter estimates, we simply substitute the necessary data from the flight of interest into (5.19) to produce a forecast of day 0 bookings. One key advantage of the full information combined model is that no missing intermediate data is required to produce a forecast. However, it should be noted that forecasts of the exogenous variables $X_{0,d-j}, X_{1,d-j}, \ldots, X_{t-1,d-j}$ are necessary.

5.5.5 Models for Passengers Boarded

In the previous section, the focus was on predicting and forecasting bookings to come for a given flight. However, the forecasting of total passengers boarded is also very important for the airlines. To be clear, total bookings are the travelers holding confirmed reservations on a flight on the departure date, while total passengers boarded are the passengers who actually board the aircraft at flight time.

Two straightforward methods exist for predicting passengers boarded. The first method is to separately predict total bookings at flight time $B_d(0)$, go-shows GS_d, no-shows NS_d, no-recs NR_d, and waitlisted passengers at flight time $WL_d(0)$. Then, given the authorized capacity of the aircraft $CAP_d(0)$, we can substitute these predicted values into equation (5.3):

$$P_d = \mathrm{MIN}[B_d(0) + WL_d(0) + GS_d + NR_d - NS_d\ ,\ CAP_d(0)] \qquad (5.20)$$

Hence, P_d becomes the predicted number of passengers boarded in class c on flight f departing on date d.

A second method of forecasting is to consider the top row of Table 5.1 as containing historical information on actual passengers boarded P_d. Then, instead of using total bookings on previous departures of the same flight number in the historical bookings model, this method uses passengers boarded on previous departures of the same flight. We replace $B_{d-j}(0)$ with P_{d-j} for all j in equation (5.14), (5.15), or (5.19) in order to develop a combined model for passengers boarded. For example, if we use equation (5.19), we obtain:

$$P_d = \sum_{j=t}^{D} \varphi_{j0} P_{d-j} + \left(\sum_{j=0}^{t-1} \varphi_{j,t-j}(P_{d-j} + g(X_{0,d-j},0,\beta) - g(X_{t-j,d-j},t-j,\beta))\right) + n^{\circ}{}_{d0} \qquad (5.21)$$

where the notation is the same as in equation (5.19). As in the previous section, we first estimate equation (5.21) via maximum likelihood estimation. Then, using advance and historical booking data from the particular flight of interest, we produce an forecast of passengers boarded. We can compare the forecasting results of equations (5.20) and (5.21) and choose the more accurate method for predicting passengers boarded.

5.6 Effect of Booking Limits on Forecasting

As presented in Chapter 2, the airline desires to maximize revenue by setting booking limits, maximum authorized limits on the number of spaces available, in each fare class on every flight. The booking limits effectively constrain the number of low fare paying travelers who make reservations early in the booking process, while setting aside spaces for high fare paying travelers who book late in the process. In setting the booking limits, the airline needs an accurate forecast of the true underlying demand -- not simply the observed demand. Thus, the forecasting methodology should take into account the effect of the booking limits on observed demand and attempt to forecast the true underlying demand.

First, we examine the case of distinct booking limits, which effectively serve as "electronic bulkheads" by dividing the cabin into fare classes. The second case is the more realistic and most commonly used scenario of nested booking limits. Nested booking limits constrain the number of spaces to be booked in lower fare classes, while setting aside spaces to be reserved in the higher fare classes. Also, in this section, we indicate how to estimate and forecast the three primary statistical models in view of the effect of the booking limits.

5.6.1 Distinct Booking Limits

Distinct booking limits basically divide the aircraft cabin into separate sections corresponding to each fare class. The sum of spaces should sum to the authorized capacity of the aircraft. In essence, the booking limits are invisible bulkheads dividing the cabin into its respective fare classes. Note that, in general, these booking limits change over time and, hence, depend on the time t days before departure. Formally, let the maximum authorized booking level (capacity) at time t days before departure for class c on flight f departing on date d be denoted by $CAP_{cfd}(t)$. Therefore, at any time t days before departure, the total bookings are limited by $CAP_{cfd}(t)$. In Chapter 2, we show a booking curve limited by a booking limit in Figure 2.5. Given that the observed bookings are restricted by the booking limits, we want to determine the true underlying demand. In the following analysis, we develop a truncated-censored approach.

Due to the finite booking limits, the data sample is an incomplete sample from the true underlying demand distribution of requests. Because of the booking limits, the airline views only the portion of the true data sample -- observed booking data which are always less than or equal to the booking limit. Thus, the data sample is said to be *censored* at $CAP_{cfd}(t)$. A censored sample is one in which some observations of the dependent variable corresponding to known sets of independent variables are not observable. [Judge et. al., 1985] If observed bookings reach the maximum level, all we know is that the underlying (unobserved) dependent variable is greater than or equal to the maximum authorized level. Empirically, we should observe a spike of demand at the booking limit (if the data is highly censored). The observable range of the dependent variable is limited to the range $[0, CAP_{cfd}(t)]$. However, the underlying (unobserved) demand can take on any value $[0, \infty)$. Figure 5.1 displays a censored Normal distribution with censoring at $CAP_{cfd}(t)$.

On the other hand, we will never have negative total bookings — observed or unobserved. This is a constraint on the distribution from which the data is drawn. A truncated distribution arises when there are restrictions on a population <u>prior to sampling</u> (Maddala, 1983). Thus, the distribution from which our sample is drawn is *truncated* at zero. No values below the lower truncation point are allowed. The probability density function, of course, must be standardized to integrate up to 1. Figure 5.2 presents a truncated Normal distribution with truncation from below at zero.

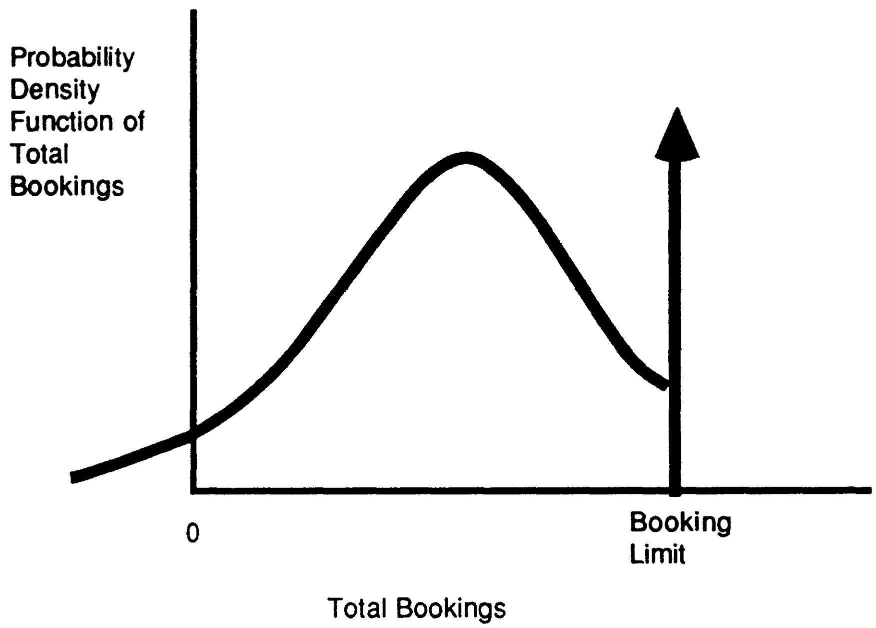

Figure 5.1 Normal Distribution Censored from above at Booking Limit

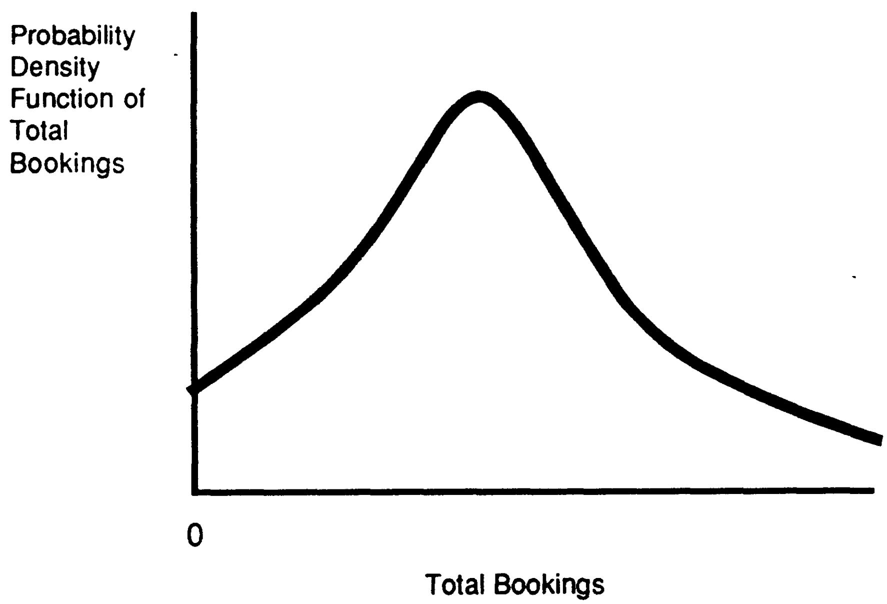

Figure 5.2 Truncated Normal Distribution from below at zero

Belobaba (1985) has extensively tested observed airline data to determine demand distribution patterns. His conclusions, slightly modified by more recent results (Lee, 1988), are that significant positive skewness of the booking distribution may occur at low levels of demand, a spike in the booking distribution may occur at high levels of demand (where booking limits have an effect), and an assumption of Normally distributed booking data seems to hold for moderate levels of demand. See Figure 5.3.

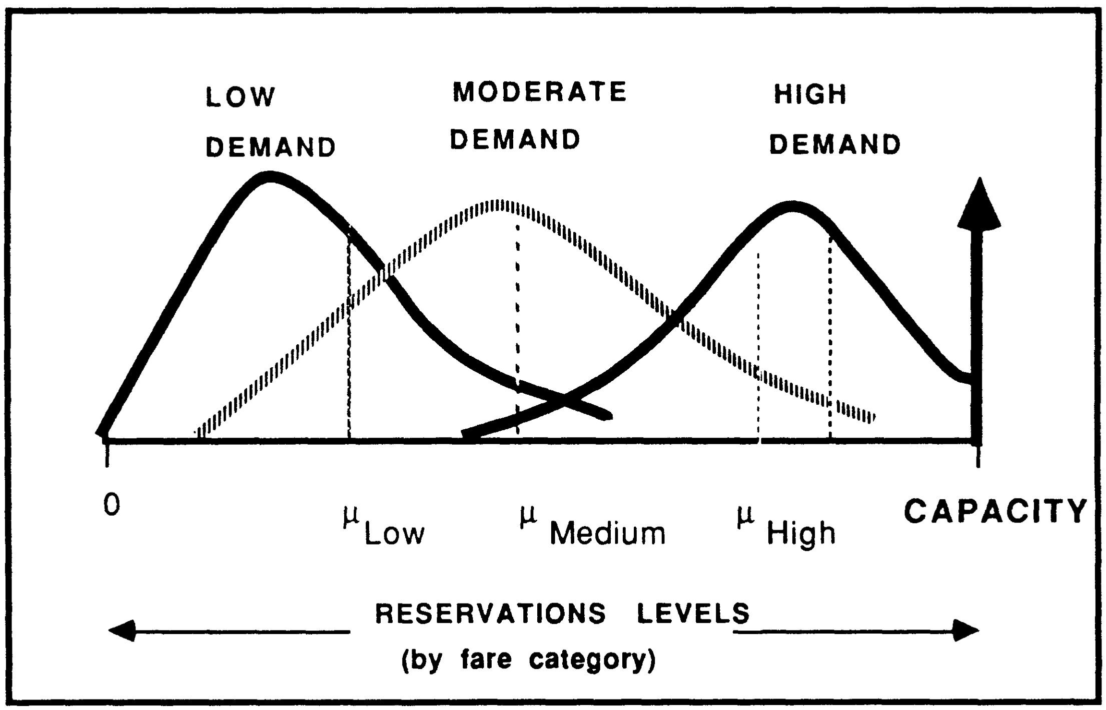

Figure 5.3 Conceptual Model of Distribution Patterns

Recasting these conclusions in terms of censored samples and truncated distributions, we develop a framework for expanded distributional analysis. The key here is to realize that truncation from below at zero may lead to the appearance of positive skewness of the distribution at low levels of demand. Likewise, censoring at authorized capacity may lead to the appearance of a spike. Finally, moderate demand levels may show very little effects of either truncation (or censoring) from above or from below -- since these limiting factors will only affect the "tails" of the distribution. Thus, this truncated-censored model of the distributional patterns is consistent with the conclusions of Belobaba (1985) and Lee (1988).

Applying the truncated–censored approach to the statistical models, we want to estimate and predict the true underlying total number of requests for bookings. Therefore, we propose three different assumptions for flights with low demand, high demand, and moderate demand. First, for flights with low demand, the truncation at zero has the most significant effect on the data. Mathematically, we have:

$$B_d(0) \geq 0$$

At worst, we will have no bookings on-hand for a particular class at flight time. For example, if the underlying distribution of requests for bookings is Normal, the total bookings on low demand flights are drawn from a truncated Normal distribution with truncation from below at zero. We assume that the fundamental relationship (5.6) holds for estimating total bookings. That is,

$$B_d(0) \;=\; B_d(t) + B_d(0,\alpha) - B_d(t,\alpha) + \varepsilon_{td}$$

Hence, the truncated Normal distribution function for low demand flights is defined as follows:

$$f(B_d(0) \mid B_d(0) \geq 0) \quad = \quad \frac{(1/\sigma)\ \phi[B_d(0) - (B_d(t) + B_d(0,\alpha) - B_d(t,\alpha))\ /\sigma]}{1 - \Phi[-\ (B_d(t) + B_d(0,\alpha) - B_d(t,\alpha))\ /\sigma]} \qquad \text{if } B_d(0) \geq 0$$

$$f(B_d(0) \mid B_d(0) \geq 0) \quad = \quad 0 \qquad \text{otherwise}$$

where $B_d(0,\alpha)$ and $B_d(t,\alpha)$ are any one of the three statistical models (advance, historical, or combined) presented in Section 5.5, ϕ is the standard Normal density function, and Φ is the cumulative standard Normal distribution function. To estimate the parameters of this model, we can form the log-likelihood function and solve for the parameters α using a maximum likelihood procedure. Finally, to produce a forecast of total bookings on the day of departure, we use the actual advance and historical booking data at time t days before departure on the particular flight of interest.

For flights with high demand, the censoring at the authorized capacity of the aircraft has the greatest effect on the data. We observe only values of total bookings less than the authorized capacity:

$$B_d(0) \leq CAP_d(0)$$

If the underlying distribution of requests for bookings is Normal, the censored Normal regression model for high demand flights is defined as follows:

$$B_d(0) \quad = \quad B_d(t) + B_d(0,\alpha) - B_d(t,\alpha) + \varepsilon_{td} \qquad \text{if RHS} < CAP_d(0)$$

$$B_d(0) \quad = \quad CAP_d(0) \qquad \text{otherwise}$$

where the error term $\varepsilon_{t0,d}$ is assumed to be Normally distributed and $B_d(0,\alpha)$ and $B_d(t,\alpha)$ are any of the three statistical models (advance, historical, or combined) described earlier in this chapter. In the econometric literature, the model is called the ***tobit model***, originally studied by Tobin (1958). As with the previous model, we can form the log-likelihood function and use a maximum likelihood procedure to estimate the parameters. To generate a forecast of total bookings on the day of departure, we use the actual advance and historical booking data at time t days before departure on the particular flight of interest.

For flights with moderate demand, the effects of the truncation and censoring may only be slight. Thus, one possible solution is to ignore these effects and estimate the parameters of the statistical models directly via ordinary least squares. A second approach is to develop a truncated–censored regression model and use maximum likelihood methods to solve for the parameters. The truncated-censored Normal regression model can be written as follows:

$$
\begin{aligned}
B_d(0) &= B_d(t) + B_d(0,\alpha) - B_d(t,\alpha) + \varepsilon_{td} && \text{if RHS} < CAP_d(0) \\
B_d(0) &= CAP_d(0) && \text{otherwise}
\end{aligned}
$$

where $B_d(0)$ is distributed as *truncated Normal with truncation from below at zero* and $B_d(0,\alpha)$ and $B_d(t,\alpha)$ are any of the three statistical models (advance, historical, or combined) for predicting airline bookings. We estimate the parameters α via maximum likelihood estimation. Then, we use the actual advance and historical booking data at time t days before departure to produce a forecast of total bookings on the day of departure for the particular flight of interest. The exact forecasting procedure depends on the particular regression model chosen for $B_d(0,\alpha)$ and $B_d(t,\alpha)$. The forecasting procedures for each regression model described earlier in the chapter are applicable for the truncated-censored model.

5.6.2 Nested Booking Limits

In most airline reservations systems, the fare classes are *nested* so that a high fare class reservation will not be denied while there are available reservations spaces. Thus, in a nested structure, we protect spaces for higher fare classes from lower fare classes. Define $SP_{cfd}(t)$ to be the minimum number of spaces protected at time t days before departure for class c on flight f departing on date d. If the number of protected spaces at any given time t days before departure is summed over all classes, then we obtain the capacity of the aircraft. Thus,

$$\sum_{c=1}^{C} SP_{cfd}(t) = TCAP_{fd}(t)$$

where $TCAP_{fd}(t)$ is the total authorized capacity of the aircraft at time t days before departure on flight f departing on date d.

In order to rigorously develop the concept of nested booking limits, let us examine the airline reservations system at such time t before any reservations have been accepted for flight f. Let f_{cfd} be the fare charged in class c on flight f departing on date d. We assume that

$$f_{1fd} > f_{2fd} > f_{3fd} > \dots > f_{Cfd}$$

Hence, the airline would profit by selling all its spaces, if possible, as class 1. So, it sets the maximum authorized booking limit, $CAP_{1fd}(t)$, for class 1 to the authorized capacity of the aircraft:

$$CAP_{1fd}(t) = TCAP_{fd}(t)$$
$$= SP_{1fd}(t) + SP_{2fd}(t) + ... + SP_{Cfd}(t)$$

Similarly, in class 2, it would profit to sell as many spaces as possible - except those set aside for class 1. Therefore, the airline sets the maximum authorized booking limit for class 2 to the authorized capacity of the aircraft less those spaces protected for class 1:

$$CAP_{2fd}(t) = TCAP_{fd}(t) - SP_{1fd}(t)$$
$$= SP_{2fd}(t) + SP_{3fd}(t) + ... + SP_{Cfd}(t)$$

For an arbitrary class i, the maximum authorized booking limit is the authorized capacity of the aircraft less those spaces set aside for classes 1,...,i - 1:

$$CAP_{ifd}(t) = TCAP_{fd}(t) - \sum_{c=1}^{i-1} SP_{cfd}(t)$$
$$= \sum_{c=i}^{C} SP_{cfd}(t)$$

This type of booking limit is called a <u>nested booking limit</u>.

As bookings begin to arrive into the airline reservations system, we need to determine the effective number of spaces available to each fare class. At any time t days before departure, the bookings for class i on flight f are limited by the maximum authorized booking limit for class i less the total number of spaces already sold in lower fare classes, less the spaces sold above the minimum number of spaces protected in higher fare class:

$$CAP_{ifd}(t) - \sum_{c=i+1}^{C} B_{cfd}(t) - \sum_{c=1}^{i-1} (MAX[0, B_{cfd}(t) - SP_{cfd}(t)]) \qquad (5.22)$$

We call (5.22) the *effective capacity* of fare class i at time t days before departure. Also, we note that the minimum number of bookings in any fare class i is zero. To develop a regression equation for the nested booking limit case, we recall the fundamental relationship for total bookings and impose the nested booking limit given in (5.22). We obtain the following:

$$B_{ifd}(0) = \quad MAX[0,MIN[F_{td}, CAP_{ifd}(t) - \sum_{c=i+1}^{C} B_{cfd}(t) - \sum_{c=1}^{i-1} (MAX[0, B_{cfd}(t) - SP_{cfd}(t)])] \qquad (5.23)$$

where $F_{td} = B_d(t) + B_d(0,\alpha) - B_d(t,\alpha) + \varepsilon_{td}$ is the expression for total bookings obtained from the fundamental relationship and $B_d(0,\alpha)$ and $B_d(t,\alpha)$ are any one of the three types of statistical models introduced earlier in this chapter. Note that equation (5.23) holds for each fare class 1,..., C.

Equation (5.23) is a truncated-censored regression model with truncation from below at zero and censoring from above at the effective capacity of fare class i. For flights which have already departed, the airline's data base contains the total bookings in each fare class, the maximum authorized capacity of each fare class, from which we can deduce the spaces protected for each fare class. Thus, we know the censoring points for all previously departed flights at any time t days before departure. We can apply maximum likelihood estimation to estimate the parameters of (5.23). Finally, we forecast total bookings on the day of departure by substituting the necessary historical and advance booking data for the particular flight of interest into (5.23). The precise forecasting procedure depends on the specific regression model chosen for $B_d(0,\alpha)$ and $B_d(t,\alpha)$. The forecasting procedures discussed earlier for each regression model are applicable in the case of nested booking limits.

5.7 Analysis of Airline Demand Distributions

In previous sections of this chapter, specific statistical models for estimating and forecasting airline demand have been introduced. However, except for illustrative purposes, we have not made specific distributional assumptions about the error terms of the statistical models. One way of analyzing the error term of the statistical models is to examine the distribution of the observed booking data. It is important to note that the probability distribution of the data will vary depending on the particular sample of booking data chosen. Thus, as emphasized throughout this thesis, it is important to select a homogeneous sample of booking data. In particular, the data sample used for distributional analysis should be categorized by flight number, fare class, day of week, and, perhaps, season of the year.

It is important to draw the distinction between *observed* airline bookings and the true, underlying airline demand pattern. Observed airline bookings are constrained by the maximum authorized booking limit. The true, underlying airline demand reflects the number of requests (whether accepted or not) received for travel. This section concentrates on possible distributions of the observed airline demand. First, we examine empirical work on airline demand distribution patterns. Second, we propose and evaluate four plausible distributions of airline bookings.

As mentioned earlier, Belobaba (1985) has empirically tested the distribution patterns of observed airline bookings (see Figure 5.3). After extensive testing, he concludes that medium demand (relative to maximum authorized capacity) flights show no significant difference from a Normal (bell shaped) distribution. However, for very low demand flights, the distribution shows significant positive skewness. For very high demand flights, the extensive computational analysis performed for the thesis shows that the distribution of the booking data has a spike at the capacity of the fare class. This evidence of a spike at capacity confirms the theoretical analysis suggested by Boeing (1982).

Since the truncation at zero and censoring at capacity have very little effect on medium demand flights, the observed data sample should be very similar to the underlying data sample. Therefore, Belobaba's results show that a Normal distribution assumption for underlying demand is quite plausible on medium demand flights. The Normal distribution, shown in Figure 5.4, is quite straightforward to use in estimation when censoring and truncation are not present. Ordinary least squares renders unbiased and efficient estimates. In the presence of significant truncation and/or censoring, it is necessary to employ maximum likelihood estimation. The primary drawback of the Normal distribution is the fact that bookings are made in discrete units, while the Normal distribution is a continuous distribution.

A second small empirical study (Brummer et. al., 1988) suggests that there may be some natural skewness in the underlying airline demand not caused by the censoring and truncation points. Thus, it proposes that underlying airline demand for bookings is Log-normally distributed. The Log-normal distribution is pictured in Figure 5.5. There are several observations about the Log-normal distribution. First, the Log-normal distribution is flexible and can take the skewness of the airline data into account. Second, the Log-normal distribution is truncated at zero. As a result, no special measures need be taken to ensure truncation at zero.

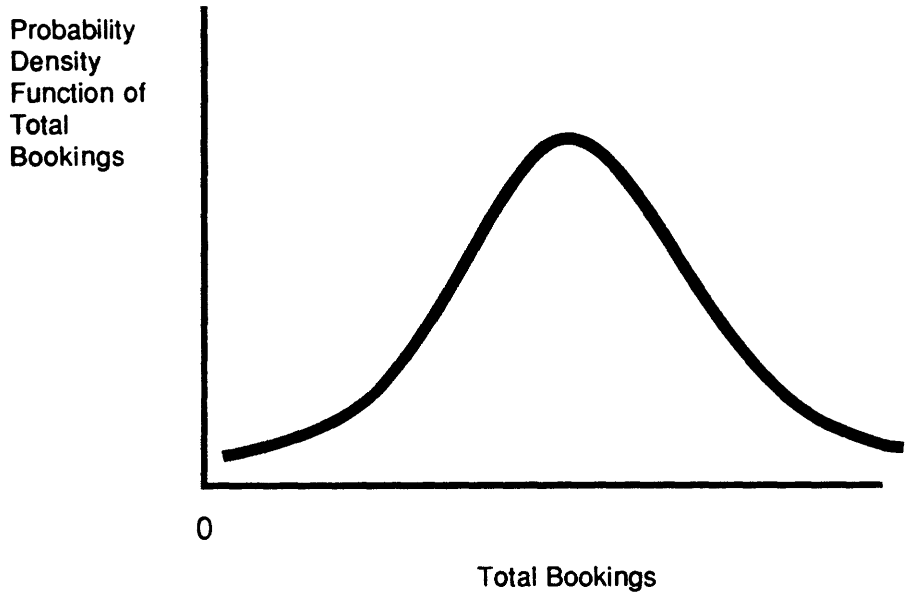

Figure 5.4 The Normal Distribution

Additionally, Log-normally distributed data can be transformed to Normally distributed data by simply taking the natural logarithm of the data. This property allows the usage of ordinary least squares methodology for parameter estimation in the non-censored cases. If the data is highly censored, then maximum likelihood estimation must be applied. It is important to note that the Log-normal distribution is only a continuous approximation to the discrete nature of airline bookings. More significantly, this distribution is not defined at zero. Since the major U.S. airlines offer 11 to 14 fare classes, it is entirely possible that one or more fare classes may not contain any bookings.

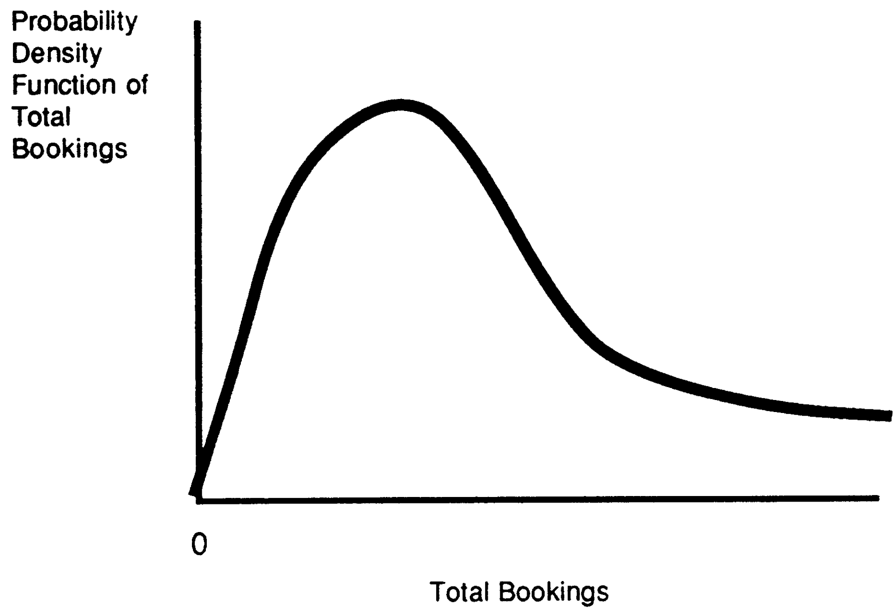

Figure 5.5 The Log-normal Distribution

In Chapter 4, the concept of the booking process as a stochastic process was introduced. After applying the Poisson approximation to the Binomial distribution in Chapter 4, total bookings are Poisson distributed with censoring at the booking limit. In effect, the booking process is modeled as a censored Poisson process with a modified request rate. Thus, the third plausible underlying distribution type is Poisson, shown in Figure 5.6. Censoring at capacity accounts for the spike on high demand flights. On low demand flights, the Poisson distribution demonstrates significant positive skewness as the mean approaches zero.

There are several important observations to be made concerning the Poisson distribution. The first observation is that it is a discrete probability distribution. Hence, it takes into account the fact that airline bookings are discrete. Second, as described in Chapter 4, the

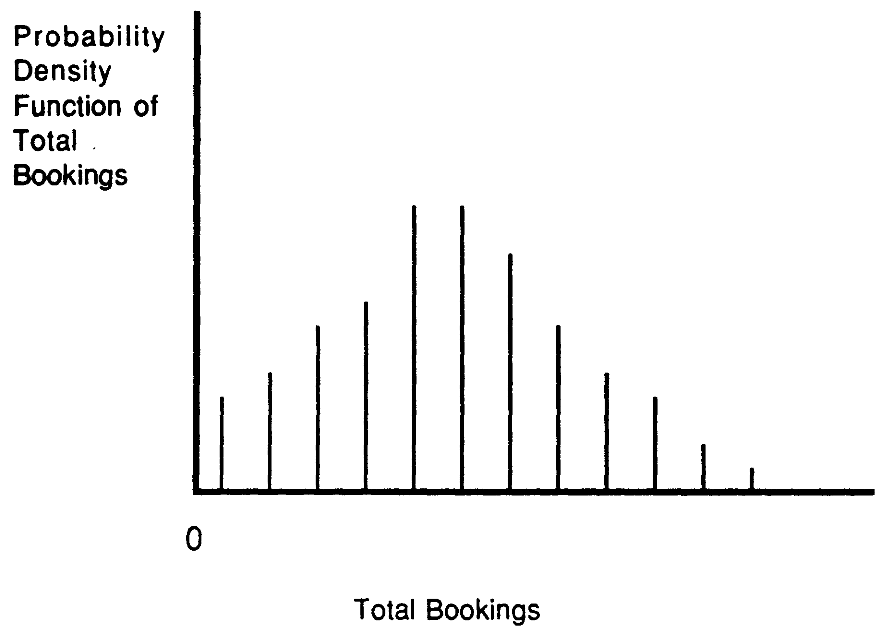

Figure 5.6 The Poisson Distribution

Poisson distribution has an intuitive probabilistic interpretation. Also, the Poisson distribution is naturally truncated at zero, never taking on negative values. Finally, the Poisson distribution assumes that the standard deviation of the demand equals the square root of the mean demand (McCullagh and Nelder, 1989). Our empirical studies show that this property of the Poisson distribution is quite reasonable.

American Airlines (Smith and Penn, 1988) suggests that the underlying demand for airline bookings is Gamma distributed. The Gamma distribution is displayed in Figure 5.7. The Gamma distributional assumption stems from the flexibility of the Gamma distribution. Whereas the Poisson distribution has only one parameter, the Gamma distribution has two parameters. A

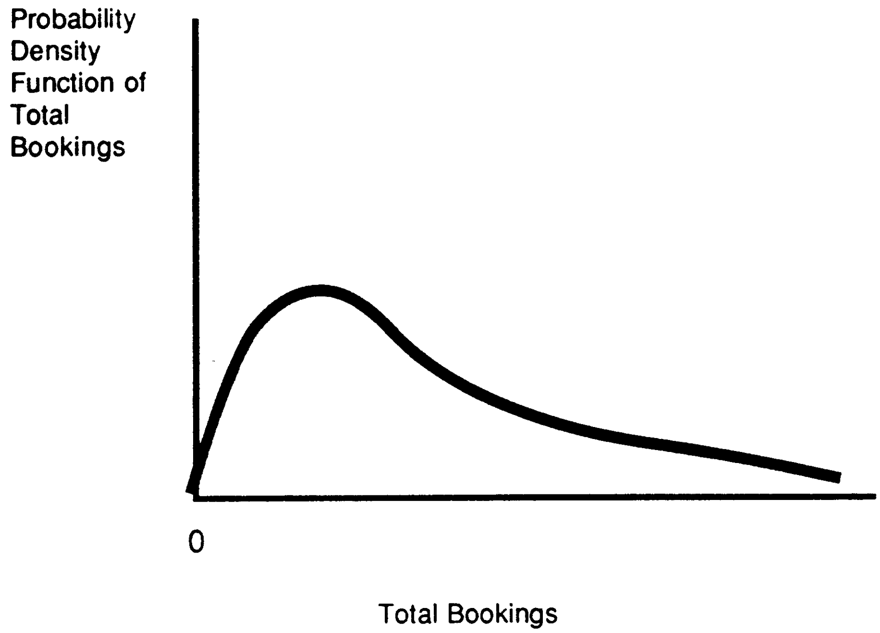

Figure 5.7 The Gamma Distribution

further benefit of the Gamma distribution is that, depending on the values of the parameters, the variance can be greater than or less than the mean of the demand. It should be noted that the Gamma distribution is a continuous distribution, which is not defined at zero. This is a drawback since bookings are inherently discrete and it is possible to have zero bookings. Finally, estimation of the Gamma distribution requires maximum likelihood estimation.

Overall, there are trade-offs between the four plausible distributional assumptions for observed bookings. The Normal distribution is computationally straightforward when significant censoring and truncation are not present. The Log-normal and Gamma distribution allow for the skewness in the data and are fairly flexible in shape. The Poisson distribution has a more natural probabilistic interpretation and takes the discrete nature of the data into account.

After considering the properties of each probability distribution, it becomes clear that the "best" choice depends on the characteristics of the particular sample of booking data under consideration. Before starting an in-depth regression analysis, it seems appropriate to do some exploratory data analysis to test the distributional patterns in the data. For example, the chi-square goodness of fit tests may be helpful in choosing a specific distribution. For details on goodness of fit tests, see Larsen and Marx (1986). Then, after fitting a particular model, visual inspection of the "residual versus fitted value" plots aid in detecting significant departures from the chosen distribution. Chapter 11 of McCullagh and Nelder (1989) describes methods for constructing these types of plots.

5.8 Conclusions

This chapter has developed a rigorous statistical framework of the airline booking process. Three types of models: advance bookings, historical bookings, and combined models have been described in detail. The combined models were the most general and intuitive models for estimation and forecasting of the booking process.

Since we want to find the distribution of requests for bookings (the true, underlying demand), we have examined the issue of booking limits and their effect on the statistical models. We have defined a measure called the effective capacity of each fare class, which takes into account the nested structure of the reservations systems of most major airlines. Our conclusion from this analysis is that the observed booking data in each class comes from a distribution which is truncated from below at zero and censored from above at the effective capacity of the class. Empirical evidence supported the conclusion of a truncated/censored model. To find the true, underlying distribution of requests for bookings, we must perform a maximum likelihood procedure to estimate the parameters. Then, we use data on-hand and the estimated parameters to forecast the true, underlying demand.

Finally, we examine several possible demand distributions for airline bookings. Ultimately, it is impossible to suggest one correct distributional assumption which will hold for every sample of airline booking data. However, the Normal distribution is the most straightforward assumption from the computational standpoint for medium demand flights which show very little effect of censoring and truncation. The Poisson distribution arises from the intuitive probabilistic model developed in Chapter 4 and has several desirable properties. However, it is computationally more intensive. In Chapter 7, we will apply the Normal and Poisson distributions to actual airline data using the models developed in this chapter.

Chapter 6 Practical Issues in Estimation and Forecasting

6.1 Introduction

Chapter 5 described a comprehensive statistical framework for estimating and forecasting the airline booking process. In this chapter, we discuss practical issues surrounding estimation and forecasting of the airline booking process. Specifically, the types of booking data generally available to airlines are described. We also outline the types of data not currently available which would aid in analyzing the airline booking process. Second, we address issues related to estimation and forecasting of airline bookings. In the estimation phase, the data itself is of most importance. The key issues to be investigated are: the amount of historical data to use in estimation, the frequency of re-estimating the models, the detection and elimination of outliers, the effect of seasonal variation, and the selection of the best model. In the forecasting phase, we will discuss how to measure the accuracy of forecasts and how to anticipate "bad" forecasts. The results of two case studies empirically show the importance of addressing the key issues in estimation and forecasting.

6.2 Airline Booking Data

This section examines the available booking data in most airline data bases. Then, we list potentially helpful additional data. For each additional type of booking data, an explanation detailing its importance for estimation and forecasting is given.

6.2.1 Available Booking Data

The airline data base for a specific fare class on a particular flight number is shown in Table 1 of Chapter 5. Each column of Table 1 contains a booking curve for a given departure date. If the flight has already departed, the corresponding column contains a *complete* booking curve with no missing data. On the other hand, if the flight has not yet departed, the column has a *partial* booking curve with data only up to the current time t days before departure. Thus, an airline data base consists of a collection of partial booking curves for flights not yet departed and complete booking curves for flights which have already departed. Depending on the airline under consideration, there may be as few as 8 weeks of complete booking curves or as many as a year or more of complete booking curves stored in the data base. Note that, due to limitations on the size of the data base, the oldest booking curve is usually deleted when a new booking curve enters the data base.

It is important to note that this data base contains only the net bookings in the reservation system at any time t before departure of a particular flight. At most airlines, no information is available on the number of requests, reservations, or cancellations made during the time before departure. Also, data on exogeneous factors which affect demand, such as major fare changes and catastrophic weather events, is very limited (if available). However, a few airlines with more sophisticated data bases may have limited data on reservations, cancellations, and exogeneous factors for each flight.

An additional type of data available to many airlines is a *seasonal index*, which attempts to account for the seasonal variation in the data. A seasonal index is a number centered around 1, which describes the deviation of any particular day from the "average" day. For example, a seasonal index of 2 on a particular day indicates that demand is usually twice as high as normal. A seasonal index of 0.5 on a specific day means that demand is generally one-half the average. The seasonal index typically is calculated on a market by market basis. Some airlines may calculate the seasonal index on a market and flight basis, creating a different seasonal index for each flight number in a particular market. In general, seasonal indices are helpful in forecasting. However, the use of a bad seasonal index can seriously hinder the estimation and forecasting of the booking process.

In this thesis, to estimate a probabilistic or statistical model for a fare class on a particular flight number, we use the historical booking curve data categorized by day of week. For example, to estimate a model for flight 33 on a Monday, we use the booking curve data from previous Mondays. Second, when specified, we use the seasonal index to take into account the seasonal variation of the data. The seasonal index is incorporated into the analysis by deseasonalizing the booking data, performing the estimation, and reseasonalizing the estimate. Alternatively, we may include the seasonal index as an explanatory variable in the statistical models.

6.2.2 Ideal Additional Booking Data

The rest of this section examines various types of booking data not generally available in airline data bases and describes their usefulness in statistical models of the booking process. First, *separate reservations and cancellation data* at each time t days before departure would be useful in estimation. This type of data allows a more direct estimation of the parameters of the probabilistic models. Furthermore, since total bookings are the difference between total reservations and total cancellations, separate reservation and cancellation data may also provide additional information for the statistical models.

The second type of desirable booking data is an *improved seasonal index*. The ideal seasonal index would be computed on a market, flight, and fare class basis, creating a separate index for each market, flight number, and fare class. The reason for a market/flight/fare class seasonal index is that each flight number and fare class combination in a market may exhibit different seasonal patterns. For example, a particular market may demonstrate high demand around a holiday period, signifying high seasonality. However, within the market, some flights may have extremely high demand and others may have lower demand. Furthermore, holiday periods traditionally demonstrate very high demand in low fare classes and very low demand in high fare classes. Thus, during a holiday period, a high fare class on a high demand flight/market combination may actually have a low seasonal index. Conversely, a low fare class on the same flight may have a very high seasonal index. Therefore, a market/flight/fare class seasonal index would be extremely helpful in predicting demand in each fare class.

The following types of additional booking data are potential explanatory variables for the statistical models. The first explanatory variable is the *percentage of travelers ticketed in each fare class*. Ticketed travelers are generally more firm about their travel plans than travelers holding unticketed reservations. Thus, this variable may help in predicting total bookings in each fare class on the day of departure. The next proposed type of booking data is the *percentage of seats sold in lower fare classes on the same flight.* This variable attempts to explain the expected amount of vertical spill from lower fare classes to higher fare classes. For example, if the lower fare classes are closed, we expect to observe higher demand than usual in the higher fare classes. A similar explanatory variable is the *percentage of seats sold in the same fare class on adjacent flights.* This variable captures the expected amount of horizontal spill from adjacent flights in the same market. If the same fare classes on adjacent flights in a market are closed, then higher demand may result on the flight under consideration. For both of the previous two variables, it would be advantageous for an airline to obtain competitors' data as well as its own data. However, competitors' data is often difficult, if not impossible, to obtain.

The next four potential explanatory variables are related to changes in the schedules and fare levels. *Changes in flight frequency in a market* from a previous scheduling period often have an effect on total bookings in each fare class. In particular, there is usually a negative correlation between increased frequency and total demand in each fare class. For example, if frequencies are added in a market, the existing demand initially tends to spread itself over all of the flights. Demand in each fare class on any particular flight decreases at first. Eventually, increased frequency may stimulate new demand. *Changes in the number of connecting flights to or from a particular flight* may affect total bookings in each fare class. Since all major U.S. airlines have large hub and spoke operations, this measure is particularly relevant. During peak months, an airline may increase the number of flights in a connecting bank of flights at a hub. The result is often higher demand on existing flights in the specific connecting bank. For off peak months, the reverse is usually true. If the number of flights in a connecting bank decreases, we often observe lower demand on the remaining flights. If an airline tends to switch aircraft types during peak and off peak periods rather than change the number of connecting flights, then a more useful measure is the changes in the number of connecting seats to or from a particular flight.

Major changes in the fare structure can dramatically affect the demand for reservations in each fare class on a flight. For example, if fares drop in a particular fare class, the result may be a large increase in demand. Conversely, if fares increase in a fare class, demand usually declines. Generally, we need to keep the date and the direction (increase or decrease) of the fare change. When estimating and forecasting on a flight leg basis in a hub and spoke environment, travelers going to many destinations are counted together. Thus, the amount of the fare change in a particular market is not very useful. More importantly, *major changes in the restrictions on any fares* can have a serious impact on the pattern of the demand in a fare class. When an advance purchase requirement is relaxed in a low fare class, there may be a shift in demand from higher fare classes to the lower fare class. Additionally, there may be a shift in when travelers book reservations. If advance purchase restrictions are loosened, travelers will book closer to the day of departure. These types of changes are very evident in the booking data that we have from a major U.S. airline.

Finally, *major non-repetitive events* have an effect on demand in particular markets. Major weather events such as Hurricane Hugo in the Southeastern U.S. and the San Francisco earthquake of 1989 can depress airline demand patterns for weeks after the event. In addition, major sports events such as the Super Bowl, the Olympics, and the World Cup may temporarily increase bookings in particular markets. If this information is kept, it would help in identifying historical outliers and temporary changes in booking patterns and, thereby, lead to increased forecasting accuracy.

6.3 Issues in Estimation and Forecasting

This section investigates key issues in the estimation and forecasting of the airline booking process. We recall the distinction made between estimation and forecasting in Section 5.4. The estimation phase fits a model to past and current data. The forecasting phase uses the estimated model to predict future, unknown values. We first discuss the important issues surrounding the estimation phase: the amount of historical data to use in the estimation procedure, the frequency of re-estimation to update the parameters of the model, how to recognize and reduce the effect of outliers, the elimination of seasonal variation, and the selection of the best model. Second, we describe the key elements of the forecasting procedure: measuring the accuracy of forecasts and methods to anticipate potential "bad" forecasts. Finally, we describe the results of two case studies, which test the effects of some of the issues on actual airline booking data.

6.3.1 The Estimation Phase

The goal of the estimation phase is to produce estimates of the parameters of the models which describe the airline booking process. In this thesis, we use the principle of maximum likelihood. The maximum likelihood estimate (MLE) of a vector of parameters β is the particular vector of values $\hat{\beta}$ which gives the greatest probability of obtaining the observed sample (Kennedy, 1987). Assuming a specific distribution of the error term, the properties of the maximum likelihood estimator are quite appealing: asymptotic unbiasedness, consistency, asymptotic efficiency. In addition, the maximum likelihood estimate is asymptotically Normally distributed. The one drawback of maximum likelihood estimation is the computational cost. However, the current generation of computer statistical software has developed relatively quick, specialized procedures for maximum likelihood estimation. We now discuss the key issues in the estimation of the airline booking process.

Amount of Data to Use in Estimation and Frequency of Re-estimation

The first issues are the amount of historical data to use in estimation and how often to re-estimate the parameters of the model. In regard to the amount of data used in estimation, econometric time series analyses often use years of data to calibrate the parameters of the model. However, as mentioned in Chapter 4, the nature of the U.S. deregulated airline environment is bringing about rapid, year to year changes. In fact, it is highly possible that any particular flight under consideration did not exist in the previous year. In addition, airline data base limitations generally prevent the use of more than one year of historical data.

Within the limitation of one year's worth of historical data, the issues remain of how much of the data to use in estimating the parameters of a model for a particular flight and how often to re-estimate the parameters. In terms of the amount of data used in estimation, it is possible to estimate a *short term* model with as little as 8 to 17 weeks of historical data or a *long term* model with as much as 26 to 52 weeks of historical data. Re-estimation of the parameters can be performed *frequently* on a weekly basis as soon as new data becomes available or *infrequently* on a monthly (or less frequent) basis after several weeks of new data has arrived.

In the computational work for this thesis, we have analyzed all four combinations: long term model with frequent re-estimation, long term model with infrequent re-estimation, short term model with frequent re-estimation, and short term model with infrequent re-estimation. Case study I later in this chapter compares a short term model with frequent re-estimation, which we call the *dynamic* estimation model, and a long term model with infrequent re-estimation, which we call the *static* estimation model.

The *dynamic* estimation model uses a small window of historical data, as little as 8 to 17 weeks of data, and re-estimates the parameters each week as the most recent historical data becomes available. The advantage of the dynamic model is that it immediately captures the short term trends in the data. However, since the dynamic model uses a small window of historical data, it may lose the stable, longer term patterns of the data. The *static* estimation model uses a larger window of historical data, from 26 to 52 weeks of data, but re-estimates the parameters only after a period of suitable length (for example, on a quarterly or semi-annual basis). The advantage of the static model is that it preserves the long term trends in the data. However, the drawback is that the short term patterns may not be captured.

Ben-Akiva and Bolduc (1987) introduce a methodology for model transferability which captures both short term and long term trends in the booking data. The goal of model transferability is to use previously estimated parameters in one portion of the data for current model estimation in the most recent portion of the data. No overlap between the two portions of data is permitted. Ben-Akiva and Bolduc propose a combined transfer estimator, expressed as a linear combination of current estimation results and previously estimated parameters. This methodology would allow us to capture the long term trends through the previously estimated parameters and the short term trends in the most recent data. Thus, we can apply the concept of model transferability by starting with estimated parameters from a year's worth of historical booking data. Then, at the end of each month, we could estimate a model solely based on the month's data. Finally, we form a combined transfer estimator by combining the previously estimated parameters and the estimated parameters from the current month's data.

The mathematical details of the combined transfer estimator are as follows. Let b_1 be the vector of previously estimated parameters and b_2 be the vector of estimated parameters on the new data. Also, let Σ_i be the covariance matrix of b_i and d be the difference between the estimators ($d = b_1 - b_2$). Then, the combined transfer estimator $\tilde{b}_2$ is given by the following equation:

$$\tilde{b}_2 = b_2 + A(b_1 - b_2)$$

where $A = \Sigma_2 (\Sigma_1 + dd' + \Sigma_2)^{-1}$ and d' denotes the transpose of d. Ben-Akiva and Bolduc point out that *the combined estimator may be viewed as an extension of the Bayesian updating procedure that explicitly accounts for the possible presence of a transfer bias.* In contrast to Bayesian updating or an estimator which uses a longer time series of data, the combined estimator does not assume that the parameters remain constant over time. Empirically, we do not test the combined estimator in this thesis. It is left as an area of further research.

A final issue concerning the amount of data to use for estimation is whether to include partial booking curve data. Many airlines currently use only complete booking curve data in their forecasting systems. The advantage is that, when booking curves are complete, an analyst can easily detect outliers and anamolies in the booking curve. On the other hand, partial booking curve data allows incorporation of the most recent trends in the data. For example, when making the transition between peak and off-peak travel periods, partial booking curve information is potentially helpful in predicting demand.

Detection and Elimination of Outliers

The second important issue in the estimation of the airline booking process is the detection and elimination of outliers. Generally, an outlier is an unusual point which has high influence on the value of the estimated parameters. In airline booking data, outliers may occur for several reasons. First, influential observations may be caused by the omission of an important causal variable. For example, a major non-recurrent event, not captured in the statistical models, may cause an unusually small number of bookings for a period of time. A second example of an omitted variable is the cancellation of a flight on another airline, which causes an increase in last minute bookings. Furthermore, a major price or fare change may initially generate unusually high or low demand. These points are outliers with respect to the historical data.

A second reason for outliers in the booking data is measurement error. In the airline industry, an airline agent is often responsible for entering the final number of passengers boarded on a flight into the computer reservations system. The agent may miscount or misenter the number of passengers boarded. If a mistake is made, it shows up in the airline data base and may appear as an outlier. A final possibility is error caused when the data is downloaded from the computer reservations system to the airline data base. In general, this error takes the form of missing values in the booking data. However, data is occasionally mismatched, which means that the wrong data is matched with the wrong flight. This mismatching can lead to the appearance of outliers.

In order to develop systematic approaches to detect outliers, it is necessary to more rigorously define a remote point. An outlier is a point which is remote in the direction of the dependent variable, remote in the direction of the explanatory variables, and whose removal causes large changes in the estimated parameters. Under the assumption of Normally distributed data, regression diagnostics aid in detecting potential outliers. To detect points remote in the direction of the dependent variable, we use the standardized residual. The standardized residual is defined as the residual divided by the standard error of the residual. If the absolute value of the standardized residual is greater than 2, then the data point is remote in the direction of the dependent variable. To identify points remote in the direction of the explanatory variables, we use the so-called "hat diagonal". The hat diagonal measures the distance between the values of the explanatory variables for a single observation and the mean of the values of the explanatory variables over the entire data set. A large hat value indicates a remote point in the direction of the explanatory variables.

Finally, to detect points whose removal causes a large change in the estimated parameters, there are several popular measures. DFFITS measures the change in the predicted value of an observation, if that observation is deleted. DFBETAS measures the change in the estimated parameters when an observation point is removed. If DFFITS or DFBETAS are larger than 2 in absolute value, then the point is potentially an outlier. Cook's D statistic and modified versions of this statistic attempt to measure the composite effect of an observation on the parameters of the model. Large values of Cook's D indicate possible outliers. This explanation of regression diagnostics constitutes only a brief overview. Detailed explanations can be found in Chatterjee and Hadi (1988) and Belsey, Kuh, and Welsch (1980). As mentioned earlier, the regression diagnostics outlined above assume Normally distributed data. However, McCullagh and Nelder (1989) extend the regression diagnostics to a broader family of models, called Generalized Linear Models. This family of models includes all of the alternative distributions of airline bookings described in Chapter 5, such as Log-normal, Poisson, and Gamma (as well as the Normal distribution). The general idea of each of the regression diagnostics remains the same, with slightly modified definitions.

In practice, airline analysts often use simple rules of thumb for detecting outliers. Two of the most frequently used approaches are the *mean and standard deviation method* and the *percentile method*. The mean and standard deviation method simply calculates the sample mean and sample standard deviation of the booking data. The criterion for identifying outliers is based on the Normal distribution property that 95% of the data should lie within two standard deviations of the mean. Thus, any points more than two standard deviations away from the mean are considered outliers. The percentile method calculates the percentiles of the data set. Any data point which falls below the 10th percentile or is above the 90th percentile is a possible outlier. Note that the percentile method guarantees that some data points will be considered outliers, while the mean and standard deviation method may not identify any data points as outliers. One of the key benefits of simple rules of thumb is that they are applied before the parameters are estimated. This saves costly computational time. The drawback is the increased risk of falsely identifying an outlier or missing an outlier altogether. As a result, valuable information may be lost from the model.

When points are identified as potential outliers, additional scrutiny should be given to the data point to see if there is some logical reason for its remoteness or influence. If no logical reason is evident, one of two methods can be applied to eliminate the effect of an outlier. First, the outlier can be *deleted* from the data set. Second, we can add a *dummy variable* to the model which equals one for the outlying data point and zero otherwise. Case studies I and II later in the chapter will demonstrate the importance of identifying and eliminating outliers from airline booking data.

Seasonal Variation

The third key element of the estimation phase is to remove seasonal variation from the booking data. Seasonal variation arises quite naturally in airline booking data. For example, travelers usually prefer sunny destinations to escape cold, winter weather. Thus, warm destinations often experience high demand during the winter months, but low demand during the summer months. Since multiple years of airline booking data may not be readily available and given that the deregulated U.S. airline environment experiences constant changes, it might not be possible to solely use booking data from previous peak seasons to predict peak demand. An airline must use some off peak booking data to predict peak demand. Hence, removal of seasonal variation in airline booking data is absolutely essential.

There are three primary methods for eliminating seasonal variation from airline booking data. We assume that a good seasonal index is available. Method 1 for removing the seasonal variation deseasonalizes all of the booking data, dividing the data by their corresponding seasonal indices. Then, we estimate the statistical model via maximum likelihood estimation. Finally, we reseasonalize the forecast of future bookings, multiplying the forecast by the seasonal index for the date of the forecast. This method attempts to eliminate all of the seasonal variation in the data before performing the estimation. The main drawback of this method is that it is not clear how to approach the booking limit in the censored models. The idea of deseasonalizing the booking limit seems illogical and counter-intuitive.

Method 2 deseasonalizes and reseasonalizes the data before estimation. To illustrate this method, suppose we have the following simple time series model:

$$B_d(0) = \alpha_1 B_{d-1}(0) + \alpha_2 B_{d-2}(0) \quad \text{for } d = 1, \ldots, D \tag{6.1}$$

where our data set consists of D observations. The method proceeds by first dividing each of the time series variables on the righthand side of (6.1) by its corresponding seasonal index. Then, we multiply each of the time series variables by the seasonal index for date d. Let SI_d be the seasonal index for date d. Equation (6.1) becomes:

$$B_d(0) = \alpha_1 \frac{SI_d}{SI_{d-1}} B_{d-1}(0) + \alpha_2 \frac{SI_d}{SI_{d-2}} B_{d-2}(0) \quad \text{for } d = 1, \ldots, D \tag{6.2}$$

The advantage of this method of removing seasonal variation is that it does not disturb the dependent variable. Thus, we need not be concerned about deseasonalizing the booking limit. In addition, the forecast does not require reseasonalization.

Method 3 for eliminating seasonal effects in airline booking data is to add a seasonal index variable to the statistical model. Since the seasonal index is centered about 1, we subtract 1 from the seasonal index before placing it in the model. The modified seasonal index has the following properties. If the index is less than 0, we expect lower than average demand. If the index is greater than 0, we expect higher than normal demand. As a result, the parameter attached to the index should be positive, measuring the effect of the seasonal index on demand. Note that this method hinges on accurate seasonal indices. Otherwise, the seasonal index variable could be quite misleading. In practice, we have found that seasonal indices vary widely in terms of accuracy. Case study I later in this section examines the effect of the removal of seasonal variation in airline booking data.

Model Selection

The final issue related to the estimation of airline booking data is model selection. In general terms, model selection is concerned with selecting a good statistical model from the set of all possible statistical models. Unfortunately, this general goal of model selection is nearly impossible to attain, given the vast array of modeling possibilities. A more limited and theoretically reasonable goal is, given a particular error distribution assumption and a set of explanatory variables, to find a model specification that fits the data well. In the statistical literature, there a number of hypothesis tests and goodness of fit indices which can used to find an acceptable model specification. These tests include the likelihood ratio test, the asymptotic t test, and the likelihood ratio index. For an in-depth explanation of these measures, see Judge et. al. (1985). In this section, we examine two topics which are particularly relevant to model selection involving airline booking data: a priori analysis and automatic model selection procedures.

The objective of a priori analysis is to develop a general idea of good model specifications before the estimation phase starts. A priori analysis is based on a combination of economic theory, knowledge of the area of application, and common sense. Generally, a priori analysis involves choosing key explanatory variables and developing an idea of the signs and magnitudes of the estimated parameters. For example, based on the knowledge of airline booking patterns, we might decide that the functional form of a booking curve for a particular fare class should be a natural logarithm. Another illustration is that, based on economic theory and common sense, the percentage of spaces sold in lower fare classes is a key explanatory variable for predicting demand in a higher fare class. Furthermore, since a higher percentage sold in the lower class brings about a higher percentage sold in the higher class, we may advance a hypothesis that the sign of the corresponding parameter should be positive. After the a priori analysis is complete, estimation of the chosen models is performed. Based on the results of the estimation, we may find an acceptable model or, on the other hand, we may need to reconsider the a priori assumptions.

Since airlines need to estimate and forecast for thousands of flights each day, the issue of automatic model selection procedures is quite important. Automatic model selection procedures are available to help determine which explanatory variables should be in the model. Montgomery and Peck (1982) describe a number of different model selection procedures. The main idea is that, using some goodness of fit criterion, the procedure iteratively selects a subset of the explanatory variables which contribute significantly to the model. Most automatic procedures are heuristic in nature. That is, there is no guarantee that the optimum model has been found, relative to the goodness of fit criterion. These heuristic procedures include *backward selection*, *forward selection*, and *stepwise regression*. Backward selection starts with all of the explanatory variables in the model and, one by one, eliminates the worst variables until all of the remaining variables contribute significantly to the model. Forward selection begins with no variables in the model and, one by one, adds significant variables to the model until no more variable contribute significantly to the model. Stepwise regression alternates between forward and backward selection until no further improvement in the model is possible.

Optimal automatic selection procedures are possible by exhaustive enumeration and tree searches. However, for even moderate numbers of explanatory variables, the running time of the procedures becomes prohibitive. Automatic procedures can produce a good model. However, automatic procedures are not a substitute for rigorous statistical analysis. A reasonable automatic selection procedure gives a listing of several possible models, each of which should be scrutinized carefully. Case study II later in this section demonstrates the effectiveness of model selection procedures for airline booking models.

6.3.2 The Forecasting Phase

The goal of the forecasting phase is to produce an intelligent prediction of future bookings, given the estimated parameters from the estimation phase and the current on-hand data from the future flight of interest. Forecasting of future airline bookings involves obtaining the estimated model from the estimation phase, substituting the appropriate values from the airline data base into the estimated model, and calculating the corresponding forecast value. We investigate two key issues in the forecasting phase: how the measure the performance of the forecasting procedure and how to anticipate "bad" forecasts.

Measuring Forecast Performance

There are several important measures of forecast accuracy which can be used to compare the forecasting ability of various models over a period of time. The first performance measure to be examined is the *mean absolute deviation* (MAD). The mean absolute deviation is the average of the absolute values of the forecast errors. Mathematically,

$$\sum_{n=1}^{N} \frac{|\text{actual}_n - \text{forecast}_n|}{N}$$

where N is the number of forecasts generated over a certain period of time. The mean absolute deviation is particularly useful when the cost of forecasting errors is proportional to the absolute size of the error. A second performance measure is the *root mean square error* (RMSE). The root mean square error is the square root of the average of the squared forecasting errors. Mathematically, we have:

$$\left(\sum_{n=1}^{N} \frac{(\text{actual}_n - \text{forecast}_n)^2}{N} \right)^{1/2}$$

where N is the number of forecasts generated over a certain period of time. Note that this measure weighs large forecast errors much more heavily than smaller errors. This measure is useful when the cost of forecasting error is proportional to the square of the error.

A third measure of forecast accuracy is the *mean absolute percentage error* (MAPE). The mean absolute percentage error is the average of the absolute values of the percentage errors. The mathematical formula for computing the MAPE over a N period forecasting horizon is:

$$\frac{1}{N} \sum_{n=1}^{N} \left| \frac{\text{actual}_n - \text{forecast}_n}{\text{actual}_n} \right| \cdot 100$$

One advantage of this measure is that it is dimensionless. A drawback for measure the accuracy of airline booking data is that the MAPE is not defined when the actual number of bookings is zero. One way to avoid this problem is to place the forecast value in the denominator, rather than the actual value. The mean absolute percentage error is particularly useful when the cost of the forecasting error is closely related to the percentage error.

When measuring forecasting performance in this thesis, we use one or more of the above performance measures. Depending on the strategy of a particular airline, one of the above measures may be more helpful than the others. For example, some airline yield management departments may take a conservative approach, desiring to minimize the possibility of large errors. In this case, the root mean square error criterion would be of interest. Another airline might feel that larger errors are acceptable when the demand is high and smaller errors are expected when the demand is low. Hence, this airline may believe that the percentage error is most relevant. The MAPE performance measure would be of primary interest. Finally, if all of the booking data is of the same order of magnitude, the mean absolute deviation may be a relevant performance measure.

In this thesis, we generate *ex post* forecasts. In essence, we split the data set into two parts. The first part corresponds to the estimation period. The second part comprises the forecasting period. For any particular forecast of future bookings, we are careful to use only the available booking data at the time that the forecast is generated. The goal is to produce forecasts using only the booking data that an airline would have available at the time of the forecast.

We use the three performance measures (MAD, MSE, and MAPE) to evaluate forecast accuracy during the forecasting period. When the data during the forecast period is not censored, we can directly compare the forecast generated and the actual number of bookings. However, when the data is censored, direct comparison is not possible since we do not observe the true, unconstrained number of bookings. Instead, Maddala (1983) suggests the calculation of the expected value of the censored random variable. Suppose that B_i^* denotes the i-th observation of the observed random variable (which is censored from above at the booking limit, CAP_i). Then, a prediction of the observed total bookings for the i-th observation is:

$$E[B_i^*] = P(B_i^* < CAP_i)^* E[B_i^* \mid B_i^* < CAP_i] + P(B_i^* = CAP_i)^* E[B_i^* \mid B_i^* = CAP_i]$$

$$= P(B_i^* < CAP_i)^* E[B_i^* \mid B_i^* < CAP_i] + P(B_i^* = CAP_i)^* CAP_i$$

For any particular distributional assumption, we can substitute the correct probability statements into the above expression. Then, we can use the resulting formula to calculate a censored forecast to compare with an actual observed value. For the censored Normal distribution, the above expression simplifies into a fairly straightforward formula (see Maddala, 1983). However, for the censored Poisson distribution, the formula is quite complex and much more difficult to apply.

Because of the complexity of the prediction formula for the censored Poisson distribution, we now develop a second method for measuring forecast performance. In most forecasting periods, there are a mixture of censored and non-censored data points. In fact, the number of non-censored points is usually far greater than the number of censored points. A reasonable way to measure forecasting accuracy is to simply remove the censored data points when computing the forecast performance measures. As a result, the possibility of misleading accuracy measures is greatly reduced. In the case studies I and II, censoring occurs only occasionally and we assume that the actual number of bookings is observed for each data point. However, in case studies III and IV in Chapter 7, we omit censored data from the calculation of the forecast performance measures.

Anticipating "Bad" Forecasts

The second key issue in the forecasting phase is how to anticipate "bad" forecasts. Forecasts are poor if they cause large forecasting errors. To be clear, it is important to distinguish between non-recurrent events which cause large forecasting errors and forecasting models that produce a bad forecast for no apparent exogeneous reason. Many non-recurrent events, such as extreme weather conditions and equipment failures, are beyond the control of the airline. Hence, large forecasting errors in these cases are not considered the result of bad forecasts. However, we want to focus on those cases of large forecasting errors where no external causes are evident. We outline two methods which may be able to anticipate some of the potentially bad forecasts.

The first method is to ensure that the values of the explanatory variables used to generate a forecast are within the range of the explanatory variables used to estimate the model. This is particularly important when a small window of data is used for estimation. Montgomery and Peck (1982) point out that *regression models are intended as interpolation equations over the range of the regressor variables used to fit the model.* Thus, before producing a forecast, we should compare the values of the explanatory variables used to generate the forecast to the explanatory variables in the data set used to estimate the model. If the forecast explanatory variables are outside the range of the explanatory variables used to estimate the model, then the forecast is suspect. A more robust estimation procedure, perhaps a moving average model or exponential smoothing, might be more appropriate in this case.

The second method of anticipating bad forecasts is to keep a running tally of the three forecast measures (MAD, RMSE, and MAPE) for each fare class on every flight number. This would allow constant monitoring of how the forecasts are performing. If the forecasting accuracy is improving or remaining constant, there is no cause for concern. However, if the forecasting accuracy is deteriorating over time, then an exception report should be generated and the model should be carefully scrutinized. In this manner, the airline can avoid a long series of "bad" forecasts by detecting a problematic model before much damage is done. This automatic monitoring of the performance of forecasting models is an important process for an airline to implement.

6.4 Case Study I: Outliers, Seasonality, and Dynamic Modeling

The main objective of this case study is to investigate the need for removal of outliers, elimination of seasonal variation, and dynamic models (which incorporate the most recent data) versus static models (which are estimated only once) in statistical models of the airline booking process. The data used in this analysis was provided by a major U.S. airline. It covers the period from October 1987 through February 1989. The data set includes four fare classes: Y, B, M, and Q, with Y being the highest fare class and Q being the lowest fare class. For each fare class, six selected markets are analyzed.

The case study starts with a static model with no outlier editing and no deseasonalization. Then, we sequentially test the effect of outlier editing, deseasonalization of the data, and dynamic modeling on the forecasting ability of the statistical models. We test forecasting ability at four time points before departure: day 7, day 14, day 21, and day 28 on data between August 1988 and February 1989. Two different statistical models are under consideration:

1. <u>8-week Moving Average Model</u>

   Bookings at day 0 = Bookings on-hand + 8-week average of bookings to come

2. Combined Regression Model

Bookings at day 0 = a_1*(Bookings on-hand) +

a_2*(8-week average of bookings to come) +

a_3*(percentage of seats sold in lower classes)

where a_i are the parameters to be estimated and the dependent variables denote total bookings in a single fare class. In this case study, the data is not highly censored. Therefore, the estimation procedure is ordinary least squares. Note that the 8-week moving average model requires no estimation. It is included in the case study to represent the airline industry standard model. Several major U.S. airlines use a model similar to the 8-week moving average model to estimate and forecast total bookings.

We decompose this case study into three steps. At each step, we compare the models which produce the most accurate forecasts for each market/class/time before departure combination. Step 1 evaluates the marginal effect of outlier editing, comparing the best statistical model without outlier editing to the best statistical model with outlier editing. Step 2 evaluates the marginal effect of deseasonalized data, given that outlier editing is performed. The second step compares the best statistical model without deseasonalized data to the best statistical model with deseasonalized data. Step 3 evaluates the marginal effect of dynamic modeling, given that outlier editing is performed and deseasonalized data is used for estimation. The third step compares the best statistical model under static modeling to the best statistical model under dynamic modeling. For each step, we report the aggregate results for each fare class and forecasting performance measure.

Step 1: The Marginal Effect of Outlier Editing

In step 1, we compare the forecast accuracy of the best model with outlier editing to the forecast accuracy of the best model without outlier editing. The mean and standard deviation method is used for outlier editing. The sample mean and standard deviation are calculated for the data. Any data point more than two standard deviations away from the mean is considered an outlier and is discarded. Table 6.1 contains the percentage improvement of the forecast accuracy of the best model with outlier editing over the forecast accuracy of the best model without outlier editing. The two measures of forecast accuracy are the root mean square error (RMSE) and mean absolute percentage error (MAPE).

CLASS	RMSE	MAPE
Y Class	8.2%	4.0%
B Class	5.3%	6.0%
M Class	19.8%	19.3%
Q Class	8.1%	3.4%

Table 6.1 Percentage Improvement of Best Model with Outlier Editing over Best Model without Outlier Editing

Table 6.1 shows that the marginal effect of outlier editing is quite significant. Outlier editing brings about a 5.3 to 19.8 percent increase in forecast accuracy under the root mean square error measure and a 3.4 to 19.3 percent rise in forecast accuracy under the mean absolute percentage error measure.

Step 2: The Marginal Effect of Deseasonalized Data

Step 2 compares the forecast accuracy of the best model with deseasonalized data to the forecast accuracy of the best model without deseasonalized data, given that outlier editing has been performed. Deseasonalized data is obtained by dividing the data by a seasonal index provided with the data set. Then, we estimate the model and produce a forecast. Finally, the forecast is reseasonalized with the seasonal index of the forecast date. The results in Table 6.2 show the percentage improvement of the forecast accuracy of the best model with deseasonalized data over the best model without deseasonalized data, given that outlier editing has been done.

CLASS	RMSE	MAPE
Y Class	2.7%	5.9%
B Class	0.5%	9.6%
M Class	4.0%	11.6%
Q Class	3.0%	11.7%

Table 6.2 Percentage Improvement of Best Model with Deseasonalized Data over Best Model without Deseasonalized Data

The marginal effect of using deseasonalized data, given that outlier editing is performed, is generally significant. The percentage improvement of the best model with deseasonalized data ranges from 0.5% to 4.0% in RMSE and from 5.9% to 11.7% in MAPE.

Step 3: The Marginal Effect of Dynamic Modeling

In step 3, we evaluate the marginal effect of dynamic modeling, given that outlier editing is done and deseasonalized data is used in the model. While the static model is estimated once at the beginning of the forecast period using the previous 32 weeks of completed booking curves, the dynamic model uses the most recent 17 week window of complete booking curves for each point in the forecasting horizon. Table 6.3 summarizes the average percentage improvement in forecast accuracy of the best dynamic model over the forecast accuracy of the best static model, given that outlier editing has been done and deseasonalized data is used.

CLASS	RMSE	MAPE
Y Class	2.4%	-7.1%
B Class	3.1%	-3.0%
M Class	-1.5%	-4.0%
Q Class	4.4%	-1.3%

Table 6.3 Percentage Improvement of Best Dynamic Model over Best Static Model

The results in Table 6.3 show that the effect of dynamic modeling is mixed. With the exception of M class, the forecasting accuracy of the best dynamic model is better than the best static model in terms of mean square error. On the other hand, the best static model clearly outperforms the best dynamic model in terms of mean absolute percentage error of forecast. When we examine the disaggregate data, the effect of dynamic modeling is still not clear. The dynamic model produces a better forecast in approximately 50% of the cases and the static model in the other half of the cases.

In conclusion, this case evaluated the marginal effects of outlier editing, deseasonalized data, and dynamic modeling. The results indicated that outlier editing significantly improves forecast accuracy and deseasonalized data (given outlier editing) generally improves forecast accuracy. The marginal effect of dynamic modeling (given outlier editing and deseasonalized data) is unclear. We note that this case study represents only a marginal analysis and further research must be done to fully evaluate the combined effects of outliers, seasonal variation, and dynamic modeling. However, we believe that this case study does show the importance of studying these issues.

6.5 Case Study II: The Effect of Outlier Editing and Model Selection

This case study tests the effect of using model selection procedures in the estimation phase on the forecasting ability of a statistical booking model. Outliers with a standardized residual greater than 2.0 in absolute value are removed from the model. The model selection procedures used in this section are the removal of explanatory variables with insignificant t-statistics and negative signs. Since all of the variables in the model shown below are booking data, we expect that none of the parameters should have a negative sign. Generally, we believe that an increase in bookings early in the process should always lead to an increase, however small, in total bookings on departure date. Similarly, under relatively homogeneous conditions, we expect that an increase in bookings on previous departures of the same flight number should generally be positively related to total bookings on the flight under consideration. Hence, total bookings on day 0 should be positively related to all of the explanatory variables.

In this section, the regression model is a combined full information model using partial booking data as well as complete booking data. The estimation procedure is ordinary least squares. In this case study, we solely forecast the period from day 28 to day 0 on each flight. Mathematically, we have:

$$\begin{aligned} \text{Day 0 Bookings} = \; & b_0{}^*(\text{SLD28}) + b_1{}^*(\text{AVG3BTC28}) + b_2{}^*(\text{AVG3BTC21}) + \\ & b_3{}^*(\text{AVG3BTC14}) + b_4{}^*(\text{AVG3BTC07}) + b_5{}^*(\text{AVGTSLD4}) + \\ & b_6{}^*(\text{AVGTSLD5}) + b_7{}^*(\text{AVGTSLD6}) + \varepsilon \end{aligned}$$

where we define SLD28 as the bookings on-hand 28 days before departure, AVG3BTC28 as the average bookings to come from day 28 to day 21 on the three most recent flights available, AVG3BTC21 as the average bookings to come from day 21 to day 14 on the three most recent flights available, AVG3BTC14 as the average bookings to come from day 14 to day 7 on the three most recent flights available, and AVG3BTC07 as the average bookings to come from day 7 to day 0 on the three most recent flights available. Note that the average bookings to come variables include any available partial booking data from flights which have not yet departed. Finally, we define AVGTSLD4 as the total booked at day 0 on the flight departing 4 weeks before departure date, AVGTSLD5 as the total booked at day 0 on the flight departing 5 weeks before departure date, AVGTSLD6 as the total booked at day 0 on the flight departing 6 weeks before departure date

First, we examine two specific flights in two major domestic markets, Denver to Minneapolis and Washington (Dulles) to Minneapolis. The first two examples produce a forecast for a single flight departure in order to specifically illustrate the mechanics of outlier editing and model selection. Then, we present general results for the Baltimore to Minneapolis market over a 7 month forecasting period.

Example 1: Denver to Minneapolis

The results presented in this example are from Q class data on a single Denver to Minneapolis flight on August 1, 1988. For estimation purposes, we use a 17 week window of deseasonalized data, consisting of the past 17 complete booking curves available on August 1, 1988. The goal is to produce a good forecast of total bookings at time t = 28 days before departure for total bookings in Q class on the Denver to Minneapolis flight departing on August 1, 1988.

First, the model is estimated and we obtain the following parameter estimates:

Variable	Value	Standardized T-statistic
b_0	0.82	1.96
b_1	2.33	1.63
b_2	1.98	1.30
b_3	1.16	2.29
b_4	-1.58	-1.64
b_5	-0.29	-1.13
b_6	0.01	0.053
b_7	0.26	1.303

The above model produces an absolute forecast error (|actual – forecast|) of 22.99 for the single flight departing on August 1, 1988.

We eliminate the explanatory variables with small t-statistics. In this case study, we take a very conservative approach and delete variables with t-statistics of less than 1.0. Thus, we eliminate the variable corresponding to b_6. Also, we note that May 1988 was an atypical month. Some special fares were introduced which affected demand in this market. Thus, we add a dummy variable which equals 1 for May 1988 observations and 0 otherwise. We obtain the following new estimates:

Variable	Value	Standardized T-statistic
b_0	0.82	2.17
b_1	2.33	1.74
b_2	1.99	1.39
b_3	1.16	2.46
b_4	-1.58	-1.76
b_5	-0.29	-1.25
b_6	------	------
b_7	0.26	1.54
Dummy	-29.94	-2.50

The absolute forecasting error with this revised model falls sharply to 6.46 for the single flight departing on August 1, 1988. The results of the estimation reveal that one of the data points has a large standardized residual of -2.09. Thus, we eliminate the data point with the large residual and re-estimate the model. We have the following results:

Variable	Value	Standardized T-statistic
b_0	0.84	1.46
b_1	4.14	2.45
b_2	2.16	1.67
b_3	1.35	2.78
b_4	-2.27	-2.65
b_5	-0.35	-1.65
b_6	------	-----
b_7	0.12	0.59
Dummy	-42.84	-3.45

The absolute forecasting error with the above model falls to 4.14 for the single flight departing on August 1, 1988. We terminated this example at this point. If we examine the above parameter estimates, we could have eliminated b_4 and b_5 because they are negative. Also, b_7 became insignificant with a t-statistic of only 0.59. However, this example certainly demonstrates the link between a well estimated model and a good forecast.

Example 2: Washington (Dulles) to Minneapolis

Example 2 presents the estimation and forecasting results for a single Washington to Minneapolis flight in Q class on December 5, 1988. A 17 week window of deseasonalized data is used for estimation purposes, consisting of the 17 most recent complete booking curves prior to December 5, 1988. We forecast total bookings at time t = 28 days before departure. This example starts by estimating the model using ordinary least squares. The estimated parameters are summarized below:

Variable	Value	Standardized T-statistic
b_0	0.71	2.91
b_1	1.17	1.12
b_2	0.44	1.25
b_3	-0.53	-0.48
b_4	-2.24	-2.00
b_5	0.17	1.19
b_6	-0.08	-0.46
b_7	0.50	3.90

This model produces a forecast with an absolute error of 5.36 for the single flight departing on December 5, 1988. Examining the results of this regression, we find that observation 5 has a large standardized residual of -2.34. We delete observation 5 and re-estimate the model. The estimated parameters are:

Variable	Value	Standardized T-statistic
b_0	0.59	2.19
b_1	1.34	1.66
b_2	0.68	2.30
b_3	-0.03	-0.03
b_4	-2.04	-2.22
b_5	0.08	0.65
b_6	-0.22	-1.65
b_7	0.68	4.69

The absolute forecasting error with this updated model decreases to 3.38 for the single flight departing on December 5, 1988. We note that two of the variables in the above estimation have t-statistics of less than 1.0. Hence, we delete b_3 and b_5 and, then, re-estimate the model. We obtain the following parameter estimates:

Variable	Value	Standardized T-statistic
b_0	0.55	3.07
b_1	1.39	2.67
b_2	0.70	2.30
b_3	------	----
b_4	-1.76	-2.98
b_5	------	----
b_6	-0.22	-3.02
b_7	0.74	7.68

The absolute forecasting error falls to 2.63 with the above model for the single flight departing on December 5, 1988. We stop the example at this point. Further investigation could be performed to eliminate the parameters with negative signs. However, the results of deleting observations with large residuals and eliminating variables with insignificant t-statistics is quite clear. The absolute forecasting error fell by roughly 50%.

Example 3: Baltimore to Minneapolis

This example compares the forecasting accuracy of the industry standard 8-week moving average model (presented in Case study I), the full information combined model with no model selection, and the full information combined model with model selection. The model selection rules are simple: eliminate variables with insignificant t-statistics and delete variables with negative estimated coefficients.

The data is from a flight in the Baltimore to Minneapolis market and is deseasonalized with a seasonal index provided by the airline. Forecasts are generated for a 7 month period from August 1988 to February 1989, using the most recent 17 weeks of complete booking curves for each forecast point. The average performance measures over the entire forecast period are given below in Table 6.4.

Model	RMSE	MAD
8-week Moving Average	11.30	8.59
Combined Regression without Model Selection	12.75	11.07
Combined Regression with Model Selection	10.09	8.16

Table 6.4 Comparison of Moving Average, Regression without Model Selection, and Regression with Model Selection

Table 6.4 shows the important of effective model selection procedures. When we compare the combined model without model selection to the combined model with model selection, there is moderate increase in forecasting accuracy. The root mean square error of forecast falls from 12.75 to 10.09 and the mean absolute deviation decreases form 11.07 to 8.16. More importantly, the combined regression model with model selection outperforms the industry standard 8-week moving average model. Note that, without model selection, the forecasting accuracy of the combined regression model is worse than the industry standard.

In this case study, we have presented three representative examples of the importance of model selection and outlier editing techniques. Throughout the extensive computational testing conducted for this thesis on other flights and fare classes, the results were similar. It is important to note that, even with model selection and outlier editing, regression models do not always outperform simple moving average models. However, in many cases, model selection and outlier editing help produce more accurate forecasts and are essential in helping regression models outperform simple moving average models.

6.6 Conclusions

This chapter has addressed the practical issues in estimation and forecasting. We show that it is necessary for estimation and forecasting methods to take into account such issues as the amount of data to use in estimation, the detection and elimination of outliers, seasonal variation, and model selection. An additional key issue is how to anticipate "bad" forecasts. Case studies I and II empirically demonstrated the importance of addressing these key issues.

The relevance of this chapter to the practical issue of using actual airline data for estimation and forecasting is unambiguous: it is necessary to put thought and effort into the estimation and forecasting process. There are key issues to deal with concerning the estimation of any statistical model. These issues must be carefully studied and analyzed in order to produce accurate forecasting models. In the next chapter, we perform model estimation and forecasting on actual airline booking data, using the censored Poisson model from Chapter 4 and the full information combined model from Chapter 5.

Chapter 7 Model Estimation and Forecasting

7.1 Introduction

Chapter 6 described the practical data issues in estimation and forecasting the airline booking process. In this chapter, we deal with the formulation and testing of probabilistic and statistical models on booking data provided by a major U.S. airline. We formulate the likelihood functions of two key probabilistic and statistical models, the censored Poisson model described in Chapter 4 and a full information combined model described in Chapter 5. Then, two case studies which apply the models to actual booking data are presented. Finally, the chapter ends with a discussion of future extensions and conclusions.

7.2 Testing of Probabilistic and Statistical Models

This section introduces two key probabilistic and statistical models, the censored Poisson model from Chapter 4 and the full information combined model from Chapter 5. For each model, we discuss how to estimate total bookings on a period-by-period basis. Then, for an arbitrary period before departure, the likelihood function and the corresponding derivatives are formulated. This section also addresses the issue of using the estimated parameters for forecasting total bookings on future flights.

7.2.1 The Censored Poisson Model

The censored Poisson model was introduced in Chapter 4 as a stochastic model of the airline booking process. This model assumes that bookings are Poisson distributed with censoring from above at capacity. Also, we pointed out that the Poisson parameter is a non-linear function of the request rate $\lambda(\tau)$ and the cancellation rate $\mu(\tau)$. We repeat the censored Poisson model given in equations (4.38) here for clarity:

$$P_n(\tau) = \exp(-m(\tau))\,\frac{m(\tau)^n}{n!}\,,\ n = 0,1,2,\ldots,CAP(\tau)-1 \tag{7.1a}$$

$$P_n(\tau) = \sum_{j=n}^{\infty} \exp(-m(\tau))\,\frac{m(\tau)^j}{j!}\,,\ n = CAP(\tau) \tag{7.1b}$$

where $m(\tau) = B(0)\exp(-\int_0^\tau \mu(\alpha)d\alpha) + \int_0^\tau \lambda(s)\exp(-\int_s^\tau \mu(\alpha)d\alpha)ds$, $P_n(\tau)$ is the probability of n bookings at time τ after the start of the booking process, and $CAP(\tau)$ is the booking limit at time τ after the start of the booking process.

In Chapter 4, we explained that it is often possible to identify subintervals of the booking process for a specific fare class on a particular flight on which the arrival rate, cancellation rate, and booking limit are approximately constant. As a result of this assumption, we showed in Chapter 4 that the likelihood function of the censored Poisson model is separable by period. Thus, the maximum likelihood estimates of λ_l and μ_l for all periods l are found as follows:

1. Identify L subintervals for each fare class on which the booking limit ($CAP(\tau) = CAP_l$) as well as the request and cancellation rates ($\lambda(\tau) = \lambda_l$ and $\mu(\tau) = \mu_l$) are approximately constant.

2. Estimate the parameters λ_l and μ_l on each of the L subintervals via maximum likelihood estimation starting with $[\tau_0, \tau_1]$ and ending with $[\tau_{L-1}, \tau_M]$, where τ_M is the time at which the flight departs.

For an arbitrary subinterval $[\tau_l, \tau_{l+1}]$, the censored Poisson distribution (7.1) becomes:

$$P_n(\tau_{l+1}) = \exp(-m(\tau_l))\,\frac{m(\tau_l)^n}{n!}\,,\ n = 0,1,2,\ldots,CAP_l-1 \qquad (7.2a)$$

$$P_n(\tau_{l+1}) = \sum_{j=n}^{\infty} \exp(-m(\tau_l))\,\frac{m(\tau_l)^j}{j!}\,,\ n = CAP_l \qquad (7.2b)$$

where $m(\tau_l) = B(\tau_l)\exp(-\mu_l(\tau_{l+1}-\tau_l)) + \frac{\lambda_l}{\mu_l}\left(1-\exp\left(-\mu_l(\tau_{l+1}-\tau_l)\right)\right)$, $P_n(\tau)$ is the probability of n bookings at time τ after the start of the booking process, and $CAP(\tau)$ is the booking limit at time t after the start of the booking process.

In order to formulate the log-likelihood function for the arbitrary interval $[\tau_l, \tau_{l+1}]$, suppose we have N observations of $(B_d(\tau_{l+1}), B_d(\tau_l))$, where d = 1, ... , N are the dates of each of the observations in the airline data base. Also, note that

$$m_d(\tau_l) = B_d(\tau_l)\exp(-\mu_l(\tau_{l+1}-\tau_l)) + \frac{\lambda_l}{\mu_l}\left(1-\exp\left(-\mu_l(\tau_{l+1}-\tau_l)\right)\right).$$

Then, suppose that we order the observations such that the first N_0 observations are observed at values less than capacity and the remaining $N - N_0$ observations are censored, observed at CAP_{ld}, where we allow the booking limit to depend on the period l and the observation d. We can formulate the likelihood function using the probabilities given in (7.2), where we let $CAP_{ld} = C_d$:

$$L^*(\lambda_l,\mu_l) = \prod_{d=1}^{N_0} \exp(-m_d(\tau_l)) \frac{m_d(\tau_l)^{B_d(\tau_l+1)}}{B_d(\tau_{l+1})!} \prod_{d=N_0+1}^{N} \sum_{j=C_d}^{\infty} \exp(-m_d(\tau_l)) \frac{m_d(\tau_l)^j}{j!}$$

The log-likelihood function $L(\lambda_l,\mu_l)$ is formed by taking the natural logarithm of both sides of the likelihood function.

$$\begin{aligned}
L(\lambda_l,\mu_l) &= \ln\left(\prod_{d=1}^{N_0} \exp(-m_d(\tau_l)) \frac{m_d(\tau_l)^{B_d(\tau_{l+1})}}{B_d(\tau_{l+1})!} \prod_{d=N_0+1}^{N} \sum_{j=C_d}^{\infty} \exp(-m_d(\tau_l)) \frac{m_d(\tau_l)^j}{j!} \right) \\
&= \ln\left(\prod_{d=1}^{N_0} \exp(-m_d(\tau_l)) \frac{m_d(\tau_l)^{B_d(\tau_{l+1})}}{B_d(\tau_{l+1})!} \right) + \ln\left(\prod_{d=N_0+1}^{N} \sum_{j=C_d}^{\infty} \exp(-m_d(\tau_l)) \frac{m_d(\tau_l)^j}{j!} \right) \\
&= \sum_{d=1}^{N_0} \ln\left(\exp(-m_d(\tau_l)) \frac{m_d(\tau_l)^{B_d(\tau_{l+1})}}{B_d(\tau_{l+1})!} \right) + \sum_{d=N_0+1}^{N} \ln\left(\sum_{j=C_d}^{\infty} \exp(-m_d(\tau_l)) \frac{m_d(\tau_l)^j}{j!} \right)
\end{aligned}$$

$$= \sum_{d=1}^{N_0} \ln\Big(\exp(-m_d(\tau_l))\Big) + \sum_{d=1}^{N_0} \ln\left(\frac{m_d(\tau_l)^{B_d(\tau_l+1)}}{B_d(\tau_{l+1})!}\right)$$

$$+ \sum_{d=N_0+1}^{N} \ln\left(\sum_{j=C_d}^{\infty} \exp(-m_d(\tau_l)) \frac{m_d(\tau_l)^j}{j!}\right)$$

$$= -\sum_{d=1}^{N_0} m_d(\tau_l) + \sum_{N_0} B_d(\tau_{l+1}) \ln(m_d(\tau_l))$$

$$- \sum_{d=1}^{N_0} \ln(B_d(\tau_{l+1})!) + \sum_{d=N_0+1}^{N} \ln\left(\sum_{j=C_d}^{\infty} \exp(-m_d(\tau_l)) \frac{m_d(\tau_l)^j}{j!}\right)$$

$$= -\sum_{d=1}^{N_0} m_d(\tau_l) + \sum_{d=1}^{N_0} B_d(\tau_{l+1}) \ln(m_d(\tau_l)) - \sum_{d=1}^{N_0} \ln(B_d(\tau_{l+1})!)$$

$$+ \sum_{d=N_0+1}^{N} \ln\left(1 - \sum_{j=0}^{C_d-1} \exp(-m_d(\tau_l)) \frac{m_d(\tau_l)^j}{j!}\right) \qquad (7.3)$$

Hence, equation (7.3) is the log-likelihood function to be maximized in order to find estimates of λ_l and μ_l on any arbitrary interval $[\tau_l, \tau_{l+1}]$. To maximize the log-likelihood function given in (7.3), we need to determine the first derivatives of $L(\lambda_l,\mu_l)$ with respect to λ_l and μ_l.

$$\frac{\delta L(\lambda_l,\mu_l)}{\delta \lambda_l} = -\sum_{d=1}^{N_0} \frac{\delta m_d(\tau_l)}{\delta \lambda_l} + \sum_{d=1}^{N_0} B_d(\tau_{l+1}) \frac{\delta \ln(m_d(\tau_l))}{\delta \lambda_l}$$

$$+ \sum_{d=N_0+1}^{N} \frac{\delta}{\delta \lambda_l} \left(\ln\left[1 - \sum_{j=0}^{C_d-1} \exp(-m_d(\tau_l)) \frac{m_d(\tau_l)^j}{j!}\right]\right)$$

$$= -\sum_{d=1}^{N_0} \frac{1}{\mu_l}(1-\exp(-\mu_l(\tau_{l+1}-\tau_l)) + \sum_{d=1}^{N_0} \frac{B_d(\tau_{l+1})}{\mu_l m_d(\tau_l)}(1-\exp(-\mu_l(\tau_{l+1}-\tau_l))))$$

$$+ \sum_{d=N_0+1}^{N} \left\{ \left[1 - \sum_{j=0}^{C_d-1} \exp(-m_d(\tau_l)) \frac{m_d(\tau_l)^j}{j!} \right]^{-1} * \frac{\delta}{\delta\lambda_l} \left(1 - \sum_{j=0}^{C_d-1} \exp(-m_d(\tau_l)) \frac{m_d(\tau_l)^j}{j!} \right) \right\}$$

(7.4)

Now, we simplify the rightmost term on the righthand side of equation (7.4).

$$\frac{\delta}{\delta\lambda_l}\left(1 - \sum_{j=0}^{C_d-1} \exp(-m_d(\tau_l)) \frac{m_d(\tau_l)^j}{j!}\right)$$

$$= -\sum_{j=0}^{C_d-1} \frac{\delta}{\delta\lambda_l}\left(\exp(-m_d(\tau_l)) \frac{m_d(\tau_l)^j}{j!}\right)$$

$$= \sum_{j=0}^{C_d-1}\left(\exp(-m_d(\tau_l)) \frac{m_d(\tau_l)^j}{j!} \frac{1}{\mu_l}(1-\exp(-\mu_l(\tau_{l+1}-\tau_l)))\right)$$

$$- \sum_{j=1}^{C_d-1}\left(\exp(-m_d(\tau_l)) \frac{m_d(\tau_l)^{j-1}}{(j-1)!} \frac{1}{\mu_l}(1-\exp(-\mu_l(\tau_{l+1}-\tau_l)))\right)$$

$$= \exp(-m_d(\tau_l)) \frac{m_d(\tau_l)^{C_d-1}}{(C_d-1)!} \frac{1}{\mu_l} (1-\exp(-\mu_l(\tau_{l+1}-\tau_l))) \qquad (7.5)$$

We substitute (7.5) into (7.4) and we find the expression for the first derivative of the log-likelihood function with respect to λ_l.

$$\frac{\delta L(\lambda_l,\mu_l)}{\delta\lambda_l} = -\sum_{d=1}^{N_0}\frac{1}{\mu_l}(1-\exp(-\mu_l(\tau_{l+1}-\tau_l)) + \sum_{d=1}^{N_0}\frac{B_d(\tau_{l+1})}{\mu_l m_d(\tau_l)}(1-\exp(-\mu_l(\tau_{l+1}-\tau_l))))$$

$$+\left\{\sum_{d=N_0+1}^{N}\left[1 - \sum_{j=0}^{C_d-1}\exp(-m_d(\tau_l))\frac{m_d(\tau_l)^j}{j!}\right]^{-1} * \right.$$

$$\left.\left(-\exp(-m_d(\tau_l))\frac{m_d(\tau_l)^{C_d-1}}{(C_d-1)!}\frac{1}{\mu_l}(1-\exp(-\mu_l(\tau_{l+1}-\tau_l)))\right)\right\} \quad (7.6)$$

The first derivative of the log-likelihood function with respect to μ_l can be derived as follows:

$$\frac{\delta L(\lambda_l,\mu_l)}{\delta\mu_l} = -\sum_{d=1}^{N_0}\frac{\delta m_d(\tau_l)}{\delta\mu_l} + \sum_{d=1}^{N_0}B_d(\tau_{l+1})\frac{\delta \ln(m_d(\tau_l))}{\delta\mu_l}$$

$$+\sum_{d=N_0+1}^{N}\frac{\delta}{\delta\mu_l}\left(\ln\left[1 - \sum_{j=0}^{C_d-1}\exp(-m_d(\tau_l))\frac{m_d(\tau_l)^j}{j!}\right]\right)$$

$$= -\sum_{d=1}^{N_0}\left[-\frac{\lambda_l}{\mu_l} + \exp(-\mu_l(\tau_{l+1}-\tau_l))\left(-(\tau_{l+1}-\tau_l)B_d(\tau_l) + \frac{\lambda_l}{\mu_l^2} + \frac{\lambda_l(\tau_{l+1}-\tau_l)}{\mu_l}\right)\right]$$

$$+\sum_{d=1}^{N_0}\frac{B_d(\tau_{l+1})}{m_d(\tau_l)}\left[-\frac{\lambda_l}{\mu_l} + \exp(-\mu_l(\tau_{l+1}-\tau_l))\left(-(\tau_{l+1}-\tau_l)B_d(\tau_l) + \frac{\lambda_l}{\mu_l^2} + \frac{\lambda_l(\tau_{l+1}-\tau_l)}{\mu_l}\right)\right]$$

$$+\sum_{d=N_0+1}^{N}\left\{\left[1 - \sum_{j=0}^{C_d-1}\exp(-m_d(\tau_l))\frac{m_d(\tau_l)^j}{j!}\right]^{-1} * \frac{\delta}{\delta\mu_l}\left(1 - \sum_{j=0}^{C_d-1}\exp(-m_d(\tau_l))\frac{m_d(\tau_l)^j}{j!}\right)\right\}$$

(7.7)

We need to simplify the rightmost term on the righthand side of equation (7.7).

$$\frac{\delta}{\delta\mu_l}\left(1 - \sum_{j=0}^{C_d-1} \exp(-m_d(\tau_l)) \frac{m_d(\tau_l)^j}{j!}\right)$$

$$= - \sum_{j=0}^{C_d-1} \frac{\delta}{\delta\lambda_l}\left(\exp(-m_d(\tau_l)) \frac{m_d(\tau_l)^j}{j!}\right)$$

$$= \sum_{j=0}^{C_d-1}\Bigg\{-\exp(-m_d(\tau_l)) \frac{m_d(\tau_l)^j}{j!}$$
$$\cdot \left[-\frac{\lambda_l}{\mu_l} + \exp(-\mu_l(\tau_{l+1}-\tau_l))\left(-(\tau_{l+1}-\tau_l)B_d(\tau_l) + \frac{\lambda_l}{\mu_l^2} + \frac{\lambda_l(\tau_{l+1}-\tau_l)}{\mu_l}\right)\right]\Bigg\}$$
$$- \sum_{j=1}^{C_d-1}\Bigg\{-\exp(-m_d(\tau_l)) \frac{m_d(\tau_l)^{j-1}}{(j-1)!}$$
$$\cdot \left[-\frac{\lambda_l}{\mu_l} + \exp(-\mu_l(\tau_{l+1}-\tau_l))\left(-(\tau_{l+1}-\tau_l)B_d(\tau_l) + \frac{\lambda_l}{\mu_l^2} + \frac{\lambda_l(\tau_{l+1}-\tau_l)}{\mu_l}\right)\right]\Bigg\}$$

$$= \exp(-m_d(\tau_l)) \frac{m_d(\tau_l)^{C_d-1}}{(C_d-1)!}$$
$$\cdot \left[-\frac{\lambda_l}{\mu_l} + \exp(-\mu_l(\tau_{l+1}-\tau_l))\left(-(\tau_{l+1}-\tau_l)B_d(\tau_l) + \frac{\lambda_l}{\mu_l^2} + \frac{\lambda_l(\tau_{l+1}-\tau_l)}{\mu_l}\right)\right] \tag{7.8}$$

Substituting equation (7.8) into (7.7), we obtain the first derivative of the log-likelihood with respect to μ_l:

$$
\begin{aligned}
\frac{\delta L(\lambda_l,\mu_l)}{\delta \mu_l} = & -\sum_{d=1}^{N_0}\left[-\frac{\lambda_l}{\mu_l}+\exp(-\mu_l(\tau_{l+1}-\tau_l))\left(-(\tau_{l+1}-\tau_l)B_d(\tau_l)+\frac{\lambda_l}{\mu_l^2}+\frac{\lambda_l(\tau_{l+1}-\tau_l)}{\mu_l}\right)\right] \\
& + \sum_{d=1}^{N_0}\frac{B_d(\tau_{l+1})}{m_d(\tau_l)}\left[-\frac{\lambda_l}{\mu_l}+\exp(-\mu_l(\tau_{l+1}-\tau_l))\left(-(\tau_{l+1}-\tau_l)B_d(\tau_l)+\frac{\lambda_l}{\mu_l^2}+\frac{\lambda_l(\tau_{l+1}-\tau_l)}{\mu_l}\right)\right] \\
& + \sum_{d=N_0+1}^{N}\left\{\left[1-\sum_{j=0}^{C_d-1}\exp(-m_d(\tau_l))\frac{m_d(\tau_l)^j}{j!}\right]^{-1} * \exp(-m_d(\tau_l))\frac{m_d(\tau_l)^{C_d-1}}{(C_d-1)!}\right. \\
& \left. * \left[-\frac{\lambda_l}{\mu_l}+\exp(-\mu_l(\tau_{l+1}-\tau_l))\left(-(\tau_{l+1}-\tau_l)B_d(\tau_l)+\frac{\lambda_l}{\mu_l^2}+\frac{\lambda_l(\tau_{l+1}-\tau_l)}{\mu_l}\right)\right]\right\}
\end{aligned}
\tag{7.9}
$$

After derivation of the log-likelihood function and the corresponding first derivatives, it is possible to maximize the log-likelihood function and solve for the maximum likelihood estimates of the parameters λ_l and μ_l. The BHHH method of maximum likelihood estimation is used for parameter estimation, since it requires only the first derivatives of the log-likelihood function (Berndt et. al., 1974).

In the censored Poisson model, a forecast of total bookings is obtained by solving for the maximum likelihood estimates on each of the L subintervals of the booking process of a given fare class on a specific flight. Let us denote the maximum likelihood estimates by $\hat{\lambda}_l$ and $\hat{\mu}_l$, for l = 1, ... , L. To find an unconstrained forecast of total bookings, we substitute the estimated parameters and the number of bookings currently on-hand into the expected value statement of the infinite capacity model, given by equation (4.44).

$$E[B(\tau_M)|B(\tau_i)] = B(\tau_i)^{*}\exp\left[\sum_{l=i}^{L-1}\hat{\mu}_l(\tau_l-\tau_{l+1})\right]$$

$$+\sum_{j=i}^{L-2}\frac{\hat{\lambda}_j}{\hat{\mu}_j}\exp\left(\sum_{k=j+1}^{L-1}\hat{\mu}_k(\tau_k-\tau_{k+1})\right)\left(1-\exp\left(\hat{\mu}_j(\tau_j-\tau_{j+1})\right)\right)$$

$$+\frac{\hat{\lambda}_{L-1}}{\hat{\mu}_{L-1}}\left(1-\exp\left(\hat{\mu}_{L-1}(\tau_{L-1}-\tau_L)\right)\right) \qquad (7.10)$$

Equation (7.10) produces a forecast of the true, underlying number of total bookings at time τ_M (the time at which the flight departs), given that there are $B(\tau_i)$ bookings on-hand at any time τ_i after the start of the booking process. Case study III later in this chapter presents results of estimation and forecasting using the censored Poisson model on actual airline data.

7.2.2 The Full Information Combined Model

The full information combined model was introduced in Chapter 5, equation (5.19) as an intuitive model of the booking process. This model expresses the booking process as a time series of historical bookings. Then, each element of the time series is viewed as the result of a booking curve. In this chapter, we focus on a special case of the full information combined model:

$$B_d(0) = \alpha_1(\text{AVG3SLD}(0)) + \alpha_2(B_d(t)) + \alpha_3(\text{AVG3BTC}(t)) + \alpha_4(\text{RATIO}(t)) + v_{d0} \qquad (7.11)$$

where $B_d(0)$ is the number of bookings at time t = 0 (the day of departure), AVG3SLD(0) is the average of day 0 bookings for the three most recent complete booking curves available for the particular flight number, $B_d(t)$ is the number of bookings on-hand at day t for the flight departing on date d, AVG3BTC(t) is the average of bookings to come from day t to day 0 on the three most recent observations available for the particular flight number, RATIO(t) is the percentage of seats sold in lower fare classes at time t on the flight departing on date d, and v_{d0} is a random error term. In this section, we assume that the dependent variable is censored Normal with censoring from above at the booking limit.

Suppose that an airline analyst is at time t = 60 days before departure of a particular flight and desires to estimate the full information combined model. Theoretically, the analyst can substitute t = 60 into equation (7.11) and estimate the model directly. However, in the extensive computational studies performed for this thesis, we have found that estimation on a interval-by-interval basis (similar to the censored Poisson model) provides more accurate forecasts. For example, the analyst might estimate equation (7.11) on the subintervals from t = 60 to t = 28 days before departure, from t = 28 to t = 14 days before departure, and from t = 14 to t = 0 days before departure. Thus, the proposed estimation method for the full information combined model is similar to the estimation method for the censored Poisson model:

1. Identify L subintervals of the forecast horizon [t,0] for each fare class on which the booking limit is approximately constant.

2. Estimate the parameters of the full information combined model on each of the L subintervals via maximum likelihood estimation starting with $[t_L, t_{L-1}]$ and ending with $[t_1, 0]$, where $t = 0$ is the time at which the flight departs.

For an arbitrary subinterval $[t_{l+1}, t_l]$, we substitute t_{l+1} for t and t_l for 0 into equation (7.11). The resulting regression equation is the following:

$$B_d(t_l) = \alpha_1(AVG3SLD(t_l)) + \alpha_2(B_d(t_{l+1})) + \alpha_3(AVG3BTC(t_{l+1})) + \alpha_4(RATIO(t_{l+1})) + v_{dt_l} \tag{7.12}$$

where $B_d(t_l)$ is the number of bookings at time $t = t_l$ days before departure, $AVG3SLD(t_l)$ is the average of day t_l bookings for the three most recent observations available for the particular flight number, $B_d(t_{l+1})$ is the number of bookings on-hand at day t_{l+1} for the flight departing on date d, $AVG3BTC(t_{l+1})$ is the average of bookings to come from time t_{l+1} to time t_l on the three most recent observations available for the particular flight number, $RATIO(t_{l+1})$ is the percentage of seats sold in lower fare classes at time t_{l+1} on the flight departing on date d , and v_{dt_l} is a random error term.

To formulate the log-likelihood function for an arbitrary interval $[t_{l+1}, t_l]$, suppose we have N observations of $B_d(t_l)$ and the corresponding booking limits $C_d(t_{l+1})$, where $d = 1, \ldots, N$ are the dates of each of the observations in the data base. As before, suppose that we order the observations such that the first N_0 of the observations are observed at values less than capacity and the remaining $N - N_0$ observations are censored from above at the booking limit, observed at $C_d(t_{l+1})$. Let us rewrite equation (7.12) more compactly as $B_d(t_l) = V_{dt_l}\alpha_{t_l} + v_{t_l}$, where V_{t_l} is the vector of the values of the explanatory variables for date d, and α_{t_l} is the vector of parameters to be estimated.

A number of authors (Maddala, 1983, Schneider, 1986, Judge et. al., 1985, and others) have formulated the log-likelihood function of a censored regression model with Normally distributed errors with mean 0 and common standard deviation σ. Thus, we summarize only the key results in this thesis. Let ϕ be the probability density function of a standard Normal random variable and Φ be the cumulative distribution function of a standard Normal random variable. Then, the likelihood function is:

$$L^*(\alpha,\sigma) = \prod_{d=1}^{N_0} \frac{1}{\sigma}\ \phi\left(\frac{B_d(t_l) - V_{dt_l}\alpha_{t_l}}{\sigma}\right) \prod_{d=N_0+1}^{N} \left[1 - \Phi\left(\frac{C_d(t_{l+1}) - V_{dt_l}\alpha_{t_l}}{\sigma}\right)\right]$$

Taking the natural logarithm of both sides of the likelihood function, we obtain the log-likelihood function $L(\alpha,\sigma)$.

$$\begin{aligned} L(\alpha,\sigma) &= \ln\left\{\prod_{d=1}^{N_0} \frac{1}{\sigma}\ \phi\left(\frac{B_d(t_l) - V_{dt_l}\alpha_{t_l}}{\sigma}\right) \prod_{d=N_0+1}^{N} \left[1 - \Phi\left(\frac{C_d(t_{l+1}) - V_{dt_l}\alpha_{t_l}}{\sigma}\right)\right]\right\} \\ &= \sum_{d=1}^{N_0} \ln\left[\frac{1}{\sigma}\ \phi\left(\frac{B_d(t_l) - V_{dt_l}\alpha_{t_l}}{\sigma}\right)\right] + \sum_{d=N_0+1}^{N} \ln\left[1 - \Phi\left(\frac{C_d(t_{l+1}) - V_{dt_l}\alpha_{t_l}}{\sigma}\right)\right] \end{aligned}$$

$$= \sum_{d=1}^{N_0} \ln\left((2\pi\sigma^2)^{-1/2}\right) - \sum_{d=1}^{N_0} \frac{1}{2}\left(\frac{(B_d(t_l) - V_{dt_l}\alpha_{t_l})^2}{\sigma^2}\right)$$

$$+ \sum_{d=N_0+1}^{N} \ln\left[1 - \Phi\left(\frac{C_d(t_{l+1}) - V_{dt_l}\alpha_{t_l}}{\sigma}\right)\right] \quad (7.13)$$

Equation (7.13) defines the log-likelihood function for the full information combined model with censored Normal errors. The first derivatives of the log-likelihood function with respect to α and σ are calculated as follows:

$$\frac{\delta L(\alpha,\sigma)}{\delta\alpha} = \sum_{d=1}^{N_0} \frac{1}{\sigma^2}\left(B_d(t_l) - V_{dt_l}\alpha_{t_l}\right)V_{dt_l}$$

$$-\frac{1}{\sigma}\sum_{d=N_0+1}^{N} V_{dt_l}\,\phi\left(\frac{C_d(t_{l+1}) - V_{dt_l}\alpha_{t_l}}{\sigma}\right)\left[1 - \Phi\left(\frac{C_d(t_{l+1}) - V_{dt_l}\alpha_{t_l}}{\sigma}\right)\right]^{-1}$$

(7.14)

$$\frac{\delta L(\alpha,\sigma)}{\delta\sigma} = -\frac{N_0}{2\sigma^2} + \sum_{d=1}^{N_0} \frac{1}{2\sigma^4}\left(B_d(t_l) - V_{dt_l}\alpha_{t_l}\right)^2$$

$$+\frac{1}{2\sigma^3}\sum_{d=N_0+1}^{N}\left(C_d(t_{l+1}) - V_{dt_l}\alpha_{t_l}\right)\phi\left(\frac{C_d(t_{l+1}) - V_{dt_l}\alpha_{t_l}}{\sigma}\right)\left[1 - \Phi\left(\frac{C_d(t_{l+1}) - V_{dt_l}\alpha_{t_l}}{\sigma}\right)\right]^{-1}$$

(7.15)

The second derivatives of the log-likelihood function can be derived, although the algebra is quite involved. Hence, we do not repeat them in this thesis. However, the reader is instructed to see Maddala (1983) for the mathematical formula for the matrix of second derivatives. With the derivation of the log-likelihood function and its first and second derivatives, the maximum likelihood estimates of the parameters α_{t_l} can be computed. The Newton-Raphson method for maximum likelihood estimation is used to calculate the parameter estimates, since first and second derivatives of the log-likelihood function are available. (See Ben-Akiva and Lerman, 1985 for a brief overview.)

In the full information combined model, a forecast of total bookings can be obtained by solving for the maximum likelihood estimates on each of the L subintervals of the booking process. Denote the maximum likelihood estimates of the parameters on the l-th subinterval by $\hat{\alpha}_{t_l}$. In order to produce an unconstrained forecast of total bookings in a particular fare class on a given flight, we use a *recursive substitution* method:

1. Start with the first interval $[t_L, t_{L-1}]$. Produce a forecast of total bookings at time t_{L-1}, $\hat{B}_d(t_{L-1})$, using equation (7.12) as follows:

   $\hat{B}_d(t_{L-1}) = \hat{\alpha}_1(\text{AVG3SLD}(t_{L-1})) + \hat{\alpha}_2(B_d(t_L)) + \hat{\alpha}_3(\text{AVG3BTC}(t_L)) + \hat{\alpha}_4(\text{RATIO}(t_L))$.

2. Proceed to the next interval $[t_{L-1}, t_{L-2}]$. Produce a forecast of total bookings at time t_{L-2}, $\hat{B}_d(t_{L-2})$, using equation (7.12) and the previous forecast of $\hat{B}_d(t_{L-1})$. Thus, we have:

   $$\hat{B}_d(t_{L-2}) = \hat{\alpha}_1(\text{AVG3SLD}(t_{L-2})) + \hat{\alpha}_2(\hat{B}_d(t_{L-1})) + \hat{\alpha}_3(\text{AVG3BTC}(t_{L-1})) + \hat{\alpha}_4(\text{RATIO}(t_{L-1})).$$

3. Repeat step 2 for successive intervals until the desired forecast of total bookings, $\hat{B}_d(0)$ is computed.

Note that a forecast of the percentage of seats sold in lower classes (RATIO) is required on each interval. If the forecasting procedure begins with the lowest fare class and works toward the highest fare class, then we will have forecasts of bookings in lower fare classes. As a result, we can produce a forecast of percentage of seats sold in lower classes. In summary, the recursive substitution method produces a forecast of the true, underlying number of total bookings. Case studies III and IV later in this chapter presents estimation and forecasting results using the full information combined model.

7.3 Case Study III: Estimation and Forecasting Results from the Censored Poisson Model and the Full Information Combined Model

Case study III tests the estimation and forecasting ability of the censored Poisson model, the full information combined model with censored Normal errors, a simple linear regression model, and an 8-week moving average model. The data set used to test this model consists of Friday departures of a single Detroit-Orlando flight between April 1989 and February 1990. The forecasting period is the last 8 weeks of data from the months of January and February 1990. The fare class examined on this flight is Q class, a class primarily booked by leisure travelers. The GAUSS software package (1988) was used to estimate the parameters of the statistical models. In this section, we examine the type of estimation performed. Then, we discuss the forecasting method. Finally, we present the results of estimation and forecasting for a specific date in the forecast horizon and summarize the results for the entire forecast period.

For each departure in the forecast period, the most recent 35 weeks of completed booking curves are included in the estimation step. We use the interval by interval estimation process. In particular, the data is split into three intervals on which the request rate, cancellation rate, and booking limits are approximately constant. The intervals of interest are: day 60 to day 28 before departure, day 28 to day 14 before departure, and day 14 to day 1 before departure. Note that we estimate total bookings on day 1 before departure rather than day 0 (the day of departure). The reliability of day 1 booking data appears to be much higher than day 0 bookings. Day 0 booking data tends to be affected by many exogeneous factors such as cancellations of flights and human error.

To forecast the unconstrained number of bookings for the censored Poisson model, we substitute the estimated parameters of each period into the expected value statement from the infinite capacity model (given in equation (7.10)). For the full information combined model, we produce unconstrained forecasts using the recursive substitution method described in Section 7.2.2. Forecasts of day 1 bookings are generated at time 60 days before departure, time 28 days before departure, and time 14 days before departure. The forecast accuracy measures, RMSE (root mean square error), MAPE (mean absolute percentage error), and MAD (mean absolute deviation), are calculated for the two month forecasting horizon.

In addition, a simple linear regression model is estimated via ordinary least squares and forecasts are generated for comparison purposes. The simple linear regression model for the interval $[\tau_l, \tau_{l+1}]$ is formulated as follows:

$$B_d(\tau_{l+1}) = \gamma_1 B_d(\tau_l) + \gamma_2 \tag{7.16}$$

where γ_1 and γ_2 are the parameters to be estimated. Forecasts for the simple linear regression model are generated by the recursive substitution method given in Section 7.2.2. Finally, the industry standard 8-week moving average model introduced in Chapter 6 is used to produce forecasts for comparison purposes.

Example of Estimation Results for January 5, 1990

Before we show the forecasting results over the entire two month forecast horizon, we demonstrate the parameter estimates of the three regression models for a single representative date during the forecast horizon. Note that no parameter estimation is required for the 8-week moving average model. For the flight departing on January 5, 1990, we estimate the censored Poisson model, the full information combined model with censored Normal errors, and the simple linear regression model using the most recent 35 weeks of complete booking data available on January 5, 1990. The results of the period by period estimations of the censored Poisson model are summarized in the following tables. The first table gives the estimated request rate per day $\hat{\lambda}$.

Censored Poisson Model	Period 1 Day 60 to 28	Period 2 Day 28 to 14	Period 3 Day 14 to 1
Request Rate $\hat{\lambda}$	0.447	0.714	0.386
Standard Error	0.091	0.796	1.480
Asymptotic t	4.910	0.898	0.261

We note that $\hat{\lambda}$ is significantly different from zero in period 1 at the 95% percent level of confidence, since the asymptotic t ratio is greater than 2. However, $\hat{\lambda}$ is not significantly different from zero in periods 2 and 3, since the asymptotic t ratios are small. We believe that the significance of the parameter estimates could be improved with the addition of data from other flights in the same market on the same day of week. Unfortunately, this data was not available.

The magnitude of the estimated request rates is as expected. The request rate is high from day 28 to day 14, which is the usual period of heavy bookings for Q class. A moderate request rate is shown in the period of day 60 to day 28 as the early booking leisure travelers make reservations. During the last two weeks, the request rate is low because advance purchase requirements preclude additional bookings in most (but not all) of the markets which flow over the flight leg under consideration.

The following table gives the estimated cancellation rate $\hat{\mu}$ per day for each of the three estimation periods:

Censored Poisson Model	Period 1 Day 60 to 28	Period 2 Day 28 to 14	Period 3 Day 14 to 1
Cancellation Rate $\hat{\mu}$	0.0	0.556	0.016
Standard Error	-----	0.326	0.040
Asymptotic t	-----	1.710	0.400

The cancellation rate $\hat{\mu}$ in period 1 approached 0 during the estimation phase. As a result, we fixed $\hat{\mu} = 0$ and estimated a censored Poisson model with no cancellations for period 1. The estimated cancellation rate for period 2 is significant at a 90% level of confidence, since the asymptotic t ratio exceeds 1.65. Finally, $\hat{\mu}$ is not significantly different from zero in period 3, since the asymptotic t ratio is very small. As in the case of the request rates, additional data could solve the problem of small asymptotic t ratios.

The magnitudes of the parameter estimates are as expected. The highest cancellation rate occurs during period 2 from day 28 to day 14 before departure. For Q class, we often observe a large number of requests and cancellations in the period two to four weeks before departure. The cancellation rate in period 3 is small, because cancellations are generally not allowed in Q class without substantial penalty after the ticket purchase deadlines.

Next, we report the results of the estimation of the full information combined model presented in equation (7.12). The following table summarizes the parameter estimates and the corresponding standard errors for each period:

Full Information Combined Model	Period 1 Day 60 to 28	Period 2 Day 28 to 14	Period 3 Day 14 to 1
AVG3SLD $\hat{\alpha}_1$	----	0.089	0.097
Standard Error	----	0.076	0.076
t-Ratio	----	1.17	1.28
$B_d(t_{t-1})$ $\hat{\alpha}_2$	----	1.20	0.853
Standard Error	----	0.12	0.097
t-Ratio	----	10.0	8.79
AVG3BTC $\hat{\alpha}_3$	0.424	----	----
Standard Error	0.111	----	----
t-Ratio	3.82	----	----
RATIO $\hat{\alpha}_4$	----	----	----
Standard Error	----	----	----
t-Ratio	----	----	----

When we estimate the full information combined model, we eliminate parameters with very small t ratios of less than 1.0 in absolute value and parameters with negative signs. Dashed lines (-----)

indicates that the corresponding variable has been eliminated from the model. We note that RATIO, the percentage of seats sold in lower fare classes, is removed from the model in all three of the periods. However, in Case study IV, we will see that the variable RATIO is indeed significant in many cases.

For the period 1 model, the only significant variable is AVG3BTC, the 3 week average of bookings to come. Since there are not many bookings on-hand at day 60, we do not expect RATIO and $B_d(t_{60})$ to be significant. For the period 2 model, we find that $B_d(t_{28})$ is significant, having a t ratio of 10.0. In particular, $B_d(t_{28})$ has a parameter slightly greater than 1.0 which suggests that increased bookings at day 28 lead to increased bookings at day 14. Finally, for the period 3 model, we observe that $B_d(t_{14})$ is significant, with a t ratio of 8.79. The parameter corresponding to $B_d(t_{14})$ is less than 1.0, reflecting the tendency of bookings in Q class to deteriorate slightly in the final days before departure.

For comparison purposes, the simple linear regression model of equation (7.16) is estimated. The parameter estimates and the standard errors for each period are given in the table below:

Simple Linear Regression Model	Period 1 Day 60 to 28	Period 2 Day 28 to 14	Period 3 Day 14 to 1
Bookings at Start of Period $\hat{\gamma}_1$	0.918	1.235	0.935
Standard Error	0.124	0.125	0.050
t-Ratio	7.40	9.88	18.70
Constant $\hat{\gamma}_2$	4.04	3.88	4.38
Standard Error	1.88	3.62	2.01
t-Ratio	2.15	1.07	2.18

The parameter estimates of $\hat{\gamma}_1$ are all significant at the 95% level of confidence, with t-ratios significantly larger than 1.96. As expected, the parameter estimates of $\hat{\gamma}_1$ are all close to 1.0. On the day 60 to day 28 subinterval, $\hat{\gamma}_1$ is slightly less than 1 which reflects the tendency of early booking travelers to change their plans. On the day 28 to day 14 subinterval, $\hat{\gamma}_1$ is slightly greater than 1 which indicates that increased bookings on-hand at day 28 leads to increased bookings at day 14. Finally, on the day 14 to day 1 subinterval, $\hat{\gamma}_1$ is slightly less than 1 which reflects the tendency of bookings on-hand in Q class to slightly deteriorate in the final days before departure.

Overall Forecasting Results

We present the aggregate forecasting performance measures over the entire 8-week forecasting horizon for all four models: the censored Poisson model (CPM), the full information combined model (FIC), the simple linear regression model (SLR), and the 8-week moving average model (MA8). The performance measures include only non-censored observations. The performance measure given include MAD (mean absolute deviation), RMSE (root mean square error), and MAPE (mean absolute percentage error). First, we summarize the forecast accuracy of the day 1 forecasts generated 60 days before departure. The following table contains the performance measures for all four models.

Accuracy Measures: Forecasts at Day 60	FIC Model	CPM Model	SLR Model	MA8 Model
MAD	4.74	5.33	10.54	6.91
MAPE	22.38	29.75	58.76	31.65
RMSE	6.40	4.42	12.43	8.89

For the day 1 forecasts produced at 60 days before departure, the full information combined model outperformed the other three models in terms of mean absolute deviation and mean absolute percentage error. However, the censored Poisson model produced the best forecasts according to the root mean square error measure. When we examined the data carefully, we found that the full information combined model produced one forecast with an extremely large error. As a result, the FIC model has a larger RMSE than the censored Poisson model. Of the three booking curve type models (CPM, SLR, and MA8), the censored Poisson model produces the most accurate forecasts by all three performance criteria.

The next table shows the results of forecasts of day 1 booking produced 28 days before departure:

Accuracy Measures: Forecasts at Day 28	FIC Model	CPM Model	SLR Model	MA8 Model
MAD	4.14	4.36	4.52	7.11
MAPE	21.26	19.62	24.36	31.91
RMSE	4.61	5.73	5.89	8.15

At day 28, the day 1 forecasts generated by the full information combined model outperform the other three models in terms of the MAD and RMSE measures. The censored Poisson model is the most accurate in terms of the MAPE performance measure. The censored Poisson model is better in MAPE because it tended to produce larger errors on high demand flights and smaller errors on low demand flights. Conversely, the full information combined model produced smaller errors on high demand flights and much larger errors on low demand flights, which leads to a higher MAPE. Of the three booking curve type models (CPM, SLR, and MA8), the censored Poisson model generates the most accurate forecasts by all three performance measures.

The final table in this case study contains the accuracy of day 1 forecasts produced at day 14 before departure by the two models:

Accuracy Measures: Forecasts at Day 14	FIC Model	CPM Model	SLR Model	MA8 Model
MAD	2.33	3.02	3.22	3.31
MAPE	10.04	15.81	13.70	14.69
RMSE	2.96	3.37	3.83	4.35

For day 1 forecasts generated at time 14 days before departure, the full information combined model produced more accurate forecasts in terms of all three performance measures. Of the three booking curve models (CPM, SLR, MA8), the censored Poisson model outperformed the other two models in terms of MAD and RMSE. However, the simple linear regression model produced the most accurate forecasts of the booking curve models in terms of MAPE.

Overall, the results of this case study are encouraging for the full information combined model. This case study clearly demonstrates that the full information combined model is a viable model for forecasting the airline booking process. Although it occasionally produced individual forecasts with large errors, the overall forecasting performance of the full information combined model is very good.

In conclusion, it is important to make two observations. First, the full information combined model is a combined booking curve/historical bookings model, while the other three models use booking curve information only. Thus, the FIC model uses more information and the resulting forecasts are expected to be more accurate. Second, in the case of the full information combined model, the estimation phase included removal of variables with insignificant t ratios and "wrong" signs, while the other three models did not include this model editing phase. This gave the full information combined model an advantage over the other three models. In view of these two observations, the censored Poisson model performed well and should be pursued in future research.

7.4 Case Study IV: Estimation and Forecasting Results from the Full Information Combined Model

Case study IV further examines the estimation and forecasting ability of the full information combined model. The error distribution assumption is censored Normal with censoring from above at the effective booking limit. In this case study, we test the full information combined model over the four different fare classes: Y, B, M, and Q. Y class is the highest fare class used primarily by business travelers, B and M are intermediate fare classes which appeal to business and some leisure travelers, and Q class is the lowest fare class used principally by leisure travelers.

The data used in this case study consists of weekly departures of a single flight number in several different markets between September 1987 and February 1989. The forecasting period is the last four months of data from November 1988 through February 1989. The SAS statistical software (1985) is used to estimate the parameters of the full information combined model. This section proceeds as follows. We first discuss the type of estimation performed and the forecasting method utilized. Then, on a fare class by fare class basis, we present the results for the entire forecast horizon.

In this case study, the full information combined model is estimated once for the entire forecast horizon. Recall the full information combined model from equation (7.12):

$$B_d(t_l) = \alpha_1(AVG3SLD(t_l)) + \alpha_2(B_d(t_{l+1})) + \alpha_3(AVG3BTC(t_{l+1})) + \alpha_4(RATIO(t_{l+1})) \qquad (7.17)$$

where $B_d(t_l)$ is the number of bookings at time $t = t_l$ days before departure, $AVG3SLD(t_l)$ is the average of day t_l bookings for the three most recent observations available for the particular flight number, $B_d(t_{l+1})$ is the number of bookings on-hand at day t_{l+1} for the flight departing on date d, $AVG3BTC(t_{l+1})$ is the average of bookings to come from time t_{l+1} to time t_l on the three most recent observations available for the particular flight number, and $RATIO(t_{l+1})$ is the percentage of seats sold in lower fare classes at time t_{l+1} on the flight departing on date d.

The data set used for estimation consists of approximately one year of weekly departures from September 1987 through October 1988. We use an interval by interval estimation process. For Q class, the intervals of interest are day 28 to 21 before departure, day 21 to 14 before departure, and day 14 to day1 before departure. For Y, B, and M classes, there are generally more reservations close to the day of departure. Therefore, the intervals of interest are day 28 to 14 before departure, day 14 to day 7 before departure, and day 7 to day 1 before departure. As in the previous case study, we estimate total bookings on day 1 because day 1 booking data is more stable than day 0 booking data.

In order to forecast the unconstrained number of bookings at day 1, we use the recursive substitution method outlined in Section 7.2.2. In essence, we take the forecast number of bookings from one interval and substitute it into the regression equation for the following interval. We continue this process recursively until a forecast of total bookings on day 1 is produced. In this case study, only forecasts of day 28 to day 1 bookings are generated. The forecast performance measures, MAD, RMSE, and MAPE, are computed for the forecasting horizon. For comparison purposes, a simple 8-week moving average model is used for forecasting. The 8-week moving average model is stated as follows:

Bookings at day 1 = Bookings On-hand at day 28 +

8-week Moving Average of Bookings to Come from day 28 to day 0

Note that the 8-week moving average term is based the past 8 available *complete* booking curves.

Example 1: Y Class, Washington (National) to Minneapolis Market

For Y Class demand on weekly Sunday departures of a single Washington to Minneapolis flight, we estimate a full information combined model with censored Normal errors. We examine the estimation results on a period by period basis. The estimated parameters for period 1, day 28 to day 14 before departure are summarized in the table below:

Period 1: Day 28 to 14	Estimate	Standard Error	Asymptotic t-Ratio
AVG3SLD	0.165	0.117	1.41
$B_d(28)$	0.876	0.097	9.03
AVG3BTC	-0.315	0.326	-0.97
RATIO	4.464	1.860	2.40

The RATIO and $B_d(28)$ estimates are significantly different from zero at the 95% level of confidence, with t-ratios greater than 1.96. The t-ratio of AVG3SLD is less than 1.65. Thus, it is not significantly different from zero at the 90% level of confidence. However, our computational experience in estimating models using airline booking data suggests that keeping any variable with a t-ratio above1.0 appears to aid in forecasting performance. Finally, the AVG3BTC variable has the wrong sign. We expect that higher bookings to come on previous flights would bring about higher total bookings on the flight under consideration. Thus, we remove AVG3BTC from the model and re-estimate the updated model:

Period 1: Day 28 to 14	Estimate	Standard Error	Asymptotic t-Ratio
AVG3SLD	0.135	0.114	1.18
$B_d(28)$	0.879	0.097	9.02
AVG3BTC	------	-----	-----
RATIO	3.634	1.674	2.17

Since all of the signs of the estimates are positive, as expected, and the asymptotic t-ratios are above 1.0, we will use the above parameter estimates for forecasting.

Next, we analyze period 2 from day 14 to day 7 before departure. The estimated parameters are as follows:

Period 2: Day 14 to 7	Estimate	Standard Error	Asymptotic t-Ratio
AVG3SLD	-0.051	0.235	-0.22
$B_d(14)$	1.169	0.222	5.27
AVG3BTC	0.187	0.346	0.54
RATIO	1.645	3.565	0.46

We notice the negative sign of the estimate of AVG3SLD. This not intuitive because we expect that higher demand on previous departures would result in higher bookings on the flight under consideration. Also, the t-ratios of AVG3BTC and RATIO are considerably less than 1.0. As a result, AVG3BTC, RATIO, and AVG3SLD are removed from the model. Note that we actually removed the variables one at a time and re-estimated the model after each variable deletion. We omit tables of the intermediate parameter estimates. The final result is that the following estimates are obtained:

Period 2: Day 14 to 7	Estimate	Standard Error	Asymptotic t-Ratio
AVG3SLD	------	------	------
$B_d(14)$	1.264	0.103	12.27
AVG3BTC	------	------	------
RATIO	------	------	------

The t-ratio of the remaining variable is quite large at 12.27. Thus, we conclude that the estimate of $B_d(14)$ is significantly different from zero.

The third period to be analyzed is day 7 to day 1 before departure. We estimate the full information combined model with censored Normal errors on period 3 data and obtain the following parameter estimates:

Period 3: Day 7 to 1	Estimate	Standard Error	Asymptotic t-Ratio
AVG3SLD	0.024	0.160	0.15
$B_d(7)$	1.075	0.092	11.68
AVG3BTC	-0.21	0.287	-0.73
RATIO	6.579	1.586	4.15

The parameter estimates associated with RATIO and $B_d(7)$ are positive and significantly different from zero at the 95% level of confidence. Conversely, AVG3BTC has the wrong sign and AVG3SLD has a very small t-ratio. Thus, we remove AVG3BTC and AVG3SLD from the model and re-estimate the parameters of the remaining variables. The updated parameter estimates are contained in the following table:

Period 3: Day 7 to 1	Estimate	Standard Error	Asymptotic t-Ratio
AVG3SLD	-----	-----	-----
$B_d(7)$	0.993	0.105	9.46
AVG3BTC	-----	-----	-----
RATIO	7.762	1.355	5.73

The remaining parameters are significantly different from zero and have positive signs.

For forecasting, the final estimates in each period are used. Recursive substitution is employed to produce the forecasts of day 1 bookings. For comparison purposes, we generate forecasts using the 8-week moving average model. The forecasting performance measures for both models over the 4 month forecasting horizon are tabulated. Also, the percentage improvement of the full information combined model over the 8-week moving average model in each performance measure is calculated.

Accuracy Measures: Day 1 Forecasts in Y Class	Full Information Combined Model	8-week Moving Average Model	Percentage Improvement
MAD	4.14	8.16	38.5%
MAPE	100.03	231.75	56.8%
RMSE	5.25	9.83	46.6%

The full information combined model with censored Normal errors outperforms the 8-week moving average model by 38.5% in mean absolute deviation, 56.8% in mean absolute percentage error, and 46.6% in root mean square error. These improvements are quite substantial, particularly for Y class demand which is usually difficult to forecast.

Example 2: B Class, Minneapolis to Baltimore Market

The data set for example 2 contains weekly, Sunday departures of a single flight number in the Minneapolis to Baltimore market. For B class, we estimate the full information combined model on three intervals: day 28 to day 14 before departure, day 14 to day 7 before departure, and day 7 to day 1 before departure. In this example, we display only the final parameter estimates after the variables with insignificant t-ratios and wrong signs have been removed. The final parameter estimates (with the asymptotic t-ratios are in parentheses) are:

Estimates (t-Ratios)	Period 1: Day 28 to 14	Period 2: Day 14 to 7	Period 3: Day 7 to 1
AVG3SLD	--------	0.096 (1.51)	0.345 (4.28)
$B_d(\cdot)$	1.188 (26.25)	1.021 (24.77)	1.031 (9.30)
AVG3BTC	0.516 (3.40)	0.417 (1.64)	------

RATIO	-------	-------	------

Note that $B_d(\cdot)$ denotes the number of bookings on-hand at the start of the period. In the above table, the variable RATIO is removed from all three periods. For B class in the Minneapolis-Baltimore market, it appears that the percentage of seats sold in lower classes may not be an important explanatory variable. On the other hand, the parameter estimates of the number of bookings on-hand at the start of the period have high t-ratios. This indicates that $B_d(\cdot)$ is a key explanatory variable.

Now, we produce forecasts of total bookings at day 1 using the above parameter estimates. The following table summarizes the forecasting performance over the 4 month horizon as well as the percentage improvement of the full information combined model over the moving average model:

Accuracy Measures: Day 1 Forecasts in B Class	Full Information Combined Model	8-week Moving Average Model	Percentage Improvement
MAD	2.727	3.50	22.1%
MAPE	124.96	163.73	23.7%
RMSE	3.291	3.954	16.8%

As in Y class, the forecasting performance of the full information combined model with censored Normal errors significantly outperforms the moving average model. The percentage improvement ranges from 16.8% in root mean square error to 23.7% in mean absolute percentage error. A combined model with censored Normal errors is potentially quite beneficial in terms of improved forecast accuracy.

Example 3: M Class, Minneapolis-San Francisco Market

Example 3 predicts day 1 bookings in M class on a weekly, Thursday flight between Minneapolis and San Francisco. We use period by period estimation, dividing the booking process into three periods. For M class, the estimation periods are day 28 to day 14, day 14 to day 7, and day 7 to day 1 before departure. As in example 2, we show only the final estimates after the variables with insignificant t-ratios and wrong signs have been deleted from the model. The final estimates (with the t-ratios in parentheses) are tabulated below:

Estimates (t-Ratios)	Period 1: Day 28 to 14	Period 2: Day 14 to 7	Period 3: Day 7 to 1
AVG3SLD	-------	0.186 (2.38)	--------
$B_d(\cdot)$	0.919 (8.69)	1.001 (9.94)	0.962 (12.70)
AVG3BTC	0.291 (1.34)	------	0.293 (1.16)
RATIO	8.376 (3.21)	2.950 (1.34)	6.193 (2.57)

$B_d(\cdot)$ denotes the number of bookings at the start of the estimation period. As before, the parameter estimates of $B_d(\cdot)$ are quite significant, indicating that this variable has high explanatory power in the model. Also, we note that the parameter corresponding to $B_d(\cdot)$ has a value close to 1.0 in most cases. Unlike B class, the percentage of seats sold in lower classes is significant in the M class model.

Using the above parameter estimates, we produce day 1 forecasts at time 28 days before departure. For comparison purposes, day 1 forecasts are also generated using the 8-week moving average model. The forecast accuracy measures are given in the following table:

Accuracy Measures: Day 1 Forecasts in M Class	Full Information Combined Model	8-week Moving Average Model	Percentage Improvement
MAD	5.573	9.615	42.0%
MAPE	19.14	34.81	45.0%
RMSE	7.233	11.41	36.6%

The full information combined model with censored Normal errors again shows significant improvement over the 8-week moving average model. The percentage improvement is 36.6% in root mean square error, 45.0% in mean absolute percentage error, and 42.0% in mean absolute error.

Example 4: Q Class, Philadelphia-Minneapolis Market

Example 4 estimates day 1 bookings in Q class on a weekly, Sunday flight between Philadelphia and Minneapolis. For Q class, we perform interval by interval estimation using the three intervals from day 28 to 21, day 21 to 14, and day 14 to 1 before departure. We estimate the full information combined model specified earlier in the chapter on each interval. The results of the final estimations, after variables with small t-ratios and wrong signs have been removed, are tabulated below:

Estimates (t-Ratios)	Period 1: Day 28 to 21	Period 2: Day 21 to 14	Period 3: Day 14 to 1
AVG3SLD	-------	-------	0.196 (1.95)
$B_d(\bullet)$	1.082 (26.11)	1.083 (40.11)	0.727 (6.03)
AVG3BTC	0.638 (3.57)	0.633 (4.61)	-------
RATIO	-------	-------	20.92 (2.14)

$B_d(\bullet)$ is the number of bookings on-hand at the beginning of the interval. In Q class, the variable RATIO is defined the percentage of seats currently sold in Q class (since there are no lower fare classes). We see that RATIO is only significant in the period immediately before departure. As in previous examples, $B_d(\cdot)$ is a highly significant variable, particularly in periods 1 and 2.

Day 1 forecasts are produced using the above parameter estimates. Also, an 8-week moving average model is used to generate day 1 forecasts for comparison purposes. The results of the forecast performance over the 4 month forecast horizon are summarized in the following table:

Accuracy Measures: Day 1 Forecasts in Q Class	Full Information Combined Model	8-week Moving Average Model	Percentage Improvement
MAD	7.257	9.009	19.5%
MAPE	31.31	37.17	18.7%
RMSE	8.795	10.872	19.1%

In Q class, the full information combined model continues to outperform the 8-week moving average model. The percentage improvement is nearly 20% in all three of the accuracy measures.

In conclusion, Case study IV demonstrates the potential benefits of a full information combined model with censored errors. In all fare classes, the combined model decisively outperforms the 8-week moving average model. The results are encouraging since the model performed well on all 4 fare classes in several different markets. Thus, the full information combined model is quite flexible. Further testing of this model is required in order to generalize the results. However, the results in this case study are quite promising.

7.5 Future Extensions

Case studies III and IV clearly demonstrate the excellent forecasting potential of the full information combined model and the censored Poisson model. In this section, we propose several extensions to the full information combined model and the censored Poisson model. First, we examine the full information combined (FIC) model. The FIC model tested in this chapter is a special case of the more general full information combined model. There are many potential combinations of explanatory variables which can be included in the model. One important extension is to use a booking curve model (instead of a 3-week moving average) to estimate the bookings to come term. Future research should address the types of functional forms for booking curves.

A second issue of further research for the full information combined model is to investigate the effect of period by period estimation for the full information combined model. The empirical results obtained in this chapter show that more intuitive parameter estimates and much more accurate forecasts result from a period by period estimation, rather than direct estimation on the entire time horizon. Future research should investigate the properties of period by period estimation for the full information combined model.

There are several potentially beneficial extensions to the censored Poisson model. First, we can extend the censored Poisson model by considering the request rate $\lambda(\tau)$ to be a function of explanatory variables, such as percentage of seats sold in lower classes and percentage of seats sold in the same fare class on adjacent flights, and so forth. In addition, the cancellation rate could also be made a function of explanatory variables, such as a variable measuring the percentage of refundable and changeable fares in a particular fare class on a specific flight.

A further issue to examine in regard to the censored Poisson model is determining the optimal number of subintervals on which to estimate the model. Our computational results indicate that it is important not to overfit a data set containing airline bookings. Forecasts produced by overfit models tend to be very poor. Further computational testing is required to investigate this issue.

A final extension is the possibility of combining the full information model and the censored Poisson model to obtain a more accurate forecasting model. There are two possibilites for combining the two models. First, we can extend the request rate $\lambda(\tau)$ of the censored Poisson model to be a function of explanatory variables. The explanatory variables can take the form of the full information combined model, thereby forming a combined model.

A second possibility is to embed the censored Poisson model within the full information model. Since the censored Poisson model is technically a booking curve model, we could substitute it directly into the general formulation of the full information combined model given in Chapter 5. The drawback of this formulation would be the complex, non-linear form of the resulting model. However, we propose a three step method for estimating this model. The first step would involve estimating the censored Poisson model on the data set under consideration. Then, we substitute the resulting estimated parameters into the bookings to come term of the general formulation of the full information combined model. The third step would be to estimate the parameters of the full information combined model. Further research is needed to ascertain the best method for combining the two models.

7.6 Conclusions

This chapter focused on the formulation and empirical testing of two probabilistic and statistical models: the censored Poisson model and the full information combined model. After we formulated the likelihood functions, the forecasting ability of the two models is demonstrated on actual airline data. In Case study III, we performed a forecasting contest between the censored Poisson model, the full information combined model, a simple linear regression model, and an 8-week moving average. The latter two models represent airline industry standard models, similar to models used at some major U.S. airlines. The full information combined model produced more accurate forecasts than the other three models in most cases. Also, given that the censored Poisson model only uses booking curve data, its forecasting performance was quite good as well. Case study IV demonstrated the excellent forecasting ability of the full information combined model over several fare classes and markets. To conclude the chapter, Section 7.5 briefly examined potential extensions of the censored Poisson model, the full information combined model, and a combined censored Poisson/full information model.

Chapter 8 Conclusions and Future Research

8.1 Introduction

The preceding chapters have explored the topic of airline reservations forecasting, from the basic definitions of the booking process and the development of a probabilistic model and a comprehensive statistical framework to the testing of the models on actual airline data. This chapter concludes this thesis by summarizing the key research findings and contributions and outlining potential future research to be done. In regard to future research, we discuss extensions to the probabilistic model and the statistical framework, extensions to include origin-destination reservations forecasting, and, finally, the need for an airline forecasting system.

8.2 Research Findings and Contributions

This thesis contributes to the area of airline reservations forecasting in several important ways. First, before performing the probabilistic and statistical analysis on actual airline data, we develop a rigorous statistical framework for analyzing the airline booking process. As noted in Chapter 3, our literature survey shows that such an analysis of the booking process has never been done previously. In the economic analysis of the booking process, we define the booking process as a series of interactions between a profit maximizing airline and a utility maximizing traveler. The results of these interactions are booking curves for each fare class on every flight on an airline. We note that the observed booking curve is constrained by the booking limit placed on each fare class. Our goal in this thesis is, thus, to forecast the true, underlying demand.

The probabilistic model in Chapter 4 views the airline booking process as a stochastic system of requests, reservations, and cancellations over the time before a flight departs. The resulting stochastic process is an immigration and death process with time-dependent request and cancellation rates. After some simplifying assumptions, we introduce a censored Poisson model for modeling the booking process. The contribution of the censored Poisson model is that, on the one hand, it captures the dynamic nature of the booking process and, on the other hand, it is not overly complex from the computational standpoint. In addition, the censored Poisson model takes into account the censoring of airline booking data from above at the booking limit.

The statistical framework examines the airline booking process from the data analysis standpoint. Three types of statistical models are introduced in this framework. The advance bookings model considers the bookings already made for a particular flight. The historical bookings model considers the bookings made on a previous departure of the same flight number. The third and most important statistical model is the combined model, which combines the advance bookings model and the historical bookings model. In particular, we develop a full information combined model, which views the booking process as a time series of historical bookings. Then, each element of the time series is viewed as a the result of a booking curve. The contribution of the full information combined model is that it intuitively combines the time series and booking curve approaches into a single model, which is shown to improve the forecasting ability over the individual models.

Additionally, the statistical framework introduces the concept of truncated-censored regression models. Airline booking data is naturally truncated at zero, since negative bookings never occur. On the other hand, booking data is censored at the booking limit. The contribution of the truncated-censored approach is that the true, underlying demand can be estimated from the observed data.

Another key contribution of this thesis is that we test the forecasting ability of the probabilistic and statistical models on actual airline data. Previous research has estimated the parameters of statistical models on actual airline data, but has not tested the forecasting performance. The extensive computational experience gained in the research for this dissertation shows that it is possible to estimate models which meet all the statistical criteria for a well estimated model but produce poor forecasts. Therefore, the results of Chapter 7 are significant because they demonstrate that the forecasts produced by the censored Poisson model and the full information combined model are more accurate than simple linear regression and moving average models in many cases.

A motivating factor for this dissertation is the simulation of the potential value of accurate forecasting to airlines discussed in Appendix A. For 2000 departures in each fare class, the simulation calculates the booking limits based on forecast demand and books reservations based on the actual demand. By varying the difference between forecast and actual demand, the simulation computes the effect of more accurate forecasts on expected revenue. The results show that more accurate forecasting is particularly valuable on high and very high demand flights. For a major U.S. airline, we calculate that, on high demand flights, each 10% of increased forecast accuracy can bring about an annual increase in revenue of 10 to 60 million dollars per year. Thus, the airline reservations forecasting framework developed in this thesis is potentially quite valuable for the airline industry.

8.2 Future Work

In this section, we discuss four of the most important directions for future research. First, we examine the extension of the stochastic model to include the case where the initial bookings in the system exceed the booking limit. The second area of further research is the extension of the statistical framework. The third area of future research is the extension of the framework to include origin-destination reservations forecasting. The final topic is the implementation of a forecasting system.

8.2.1 Extension of the General Probabilistic Model

The single fare class, finite capacity probabilistic model developed in Chapter 4 (equation (4.29)) assumes that the initial number of bookings in the booking process is less than the booking limit. Since the booking limit changes over time, it is important to understand that total bookings may occasionally exceed the booking limit. For example, if an airline seat inventory control analyst decides that "too many" seats have been sold in a particular fare class, the analyst may set the booking limit below the current total bookings so that no further bookings are allowed. In fact, the key observation is that *only cancellations* are allowed in the booking process until the total bookings on-hand fall below the booking limit. Therefore, the single fare class, finite capacity model should be extended to allow cancellations when the number of bookings exceeds the booking limit.

We re-formulate the single fare class, finite capacity model to allow cancellations when total bookings exceed the booking limit. At any time τ days after the booking process begins, the booking limit of fare class c, CAP(τ), is known and fixed. When total bookings in fare class c are less than the booking limit, reservations and cancellations are allowed. When total bookings exceed the booking limit, only cancellations are allowed. Thus, the state space of the stochastic process is constrained to the set of non-negative integers less than or equal to the maximum of B(τ) and CAP(τ).

As in Chapter 4, B(τ) is the number of bookings in fare class c on flight f departing on date d on-hand at time τ days after the airline has started accepting reservations. Requests for reservations are assumed to arrive in a Poisson manner with time-dependent rate $\lambda(\tau)$, $0 \le \tau \le M$, independent of population size. However, because of the booking limit, the request rate $\lambda(\tau)$ falls to 0 when the number of bookings is greater than CAP(τ). Cancellations occur in a random manner, where $\mu(\tau)$ is the time-dependent cancellation rate.

Now, we are able to write the conditional probabilities for the finite capacity case. We assume that the period $\Delta\tau$ is very short, so that at most one request, one cancellation, or nothing at all occurs. The conditional probabilities are:

$$P[B(\tau+\Delta\tau) = n+1 \mid B(\tau) = n] = \begin{cases} \lambda(\tau)\Delta\tau + o_1(\Delta\tau), & n = 0,1,\ldots,CAP(\tau)-1 \\ 0 & \text{otherwise} \end{cases} \tag{8.1a}$$

$$P[B(\tau+\Delta\tau) = n-1 \mid B(\tau) = n] = \begin{cases} n\mu(\tau)\Delta\tau + o_2(\Delta\tau), & n = 0,1,\ldots, \max(CAP(\tau), B(\tau)) \\ 0 & \text{otherwise} \end{cases} \tag{8.1b}$$

$$P[B(\tau+\Delta\tau) = n \mid B(\tau) = n] = \begin{cases} 1 - \lambda(\tau)\Delta\tau - n\mu(\tau)\Delta\tau + o_3(\Delta\tau), & n = 0,1,\ldots,CAP(\tau)-1 \\ 1 - n\mu(\tau)\Delta\tau + o_3(\Delta\tau), & n = CAP(\tau),\ldots,\max(CAP(\tau),B(\tau)) \\ 0 \text{ otherwise} \end{cases}$$

(8.1c)

$$P[B(\tau+\Delta\tau) = k \mid B(\tau) = n] = \begin{cases} o_4(\Delta\tau) & \text{for } \mid k - n \mid > 1 \\ 0 \text{ otherwise} \end{cases}$$

(8.1d)

where $o_i(\Delta\tau)$ represents higher order terms such that $\lim_{\Delta\tau\to 0}\left[\frac{o_i(\Delta\tau)}{\Delta\tau}\right] = 0$ and $\sum_{i=1}^{4} o_i(\Delta\tau) = 0$.

If we let $P_n(\tau) = P[B(\tau) = n]$, then the differential equations describing the airline booking process for $n = 0,1,\ldots,CAP(\tau)-1$ are the same as in case 1. For $n = 0,1,\ldots,CAP(\tau)-1$, equations (8.1) reduce to equations (4.1) and, hence, the differential equation (4.4) still holds:

$$\frac{dP_n(\tau)}{d\tau} = -(\lambda(\tau) + n\mu(\tau))P_n(\tau) + (n+1)\mu(\tau)P_{n+1}(\tau) + \lambda(\tau)P_{n-1}(\tau),\ 0 \le n \le CAP(\tau)-1 \qquad (8.2)$$

For $n = CAP(\tau)$, we can write the following equation (ignoring higher order terms):

$$P_n(\tau+\Delta\tau) = P_{n+1}(\tau)(n+1)\mu(\tau)\Delta\tau + P_n(\tau)(1-(n+1)\mu(\tau)\Delta\tau) + \lambda(\tau)P_{n-1}(\tau)\Delta\tau$$

This equation holds true since, for $n = CAP(\tau)$, the state involving exactly n bookings in the interval $[0, \tau+\Delta\tau]$ is obtained from n+1 bookings in $[0,\tau]$ with 1 cancellation in time $\Delta\tau$ *or* from n bookings in $[0,\tau]$ and nothing happening in time $\Delta\tau$ *or* from n–1 bookings in $[0,\tau]$ and one request in time $\Delta\tau$.

We rearrange terms and divide by $\Delta\tau$:

$$\frac{P_n(\tau+\Delta\tau) - P_n(\tau)}{\Delta\tau} = P_{n+1}(\tau)(n+1)\mu(\tau) - P_n(\tau)(n-1)\mu(\tau) + \lambda(\tau)P_{n-1}(\tau)$$

As $\Delta\tau \to 0$, we have the following differential equation for $n = CAP(\tau)$:

$$\frac{dP_n(\tau)}{d\tau} = P_{n+1}(\tau)(n+1)\mu(\tau) - P_n(\tau)(n-1)\mu(\tau) + \lambda(\tau)P_{n-1}(\tau) \tag{8.3}$$

Finally, for $n = CAP(\tau) + 1, \ldots, \max(B(\tau), CAP(\tau))$, we can write the following equation (ignoring higher order terms):

$$P_n(\tau+\Delta\tau) = P_{n+1}(\tau)(n+1)\mu(\tau)\Delta\tau + P_n(\tau)(1-(n+1)\mu(\tau)\Delta\tau)$$

This equation holds true since, for $n = CAP(\tau) + 1, \ldots, \max(B(\tau), CAP(\tau))$, the state involving exactly n bookings in the interval $[0, \tau+\Delta\tau]$ is obtained from n+1 bookings in $[0,\tau]$ with 1 cancellation in time $\Delta\tau$ *or* from n bookings in $[0,\tau]$ and nothing happening in time $\Delta\tau$. That is, $n = CAP(\tau) + 1, \ldots, \max(B(\tau), CAP(\tau))$, reservations are not allowed. We rearrange terms and divide by $\Delta\tau$:

$$\frac{P_n(\tau+\Delta\tau) - P_n(\tau)}{\Delta\tau} = P_{n+1}(\tau)(n+1)\mu(\tau) - P_n(\tau)(n-1)\mu(\tau)$$

As $\Delta\tau \to 0$, we have the following differential equation for $n = CAP(\tau) + 1, \ldots, \max(B(\tau), CAP(\tau))$:

$$\frac{dP_n(\tau)}{d\tau} = P_{n+1}(\tau)(n+1)\mu(\tau) - P_n(\tau)(n-1)\mu(\tau) \tag{8.4}$$

The initial condition is $P_m(0) = \begin{cases} 1 & \text{if } m = B(0) \\ 0 & \text{otherwise} \end{cases}$ since we start with B(0) bookings in the system at time 0 of the booking process.

We want to solve the system of differential equations formed by (8.2), (8.3), and (8.4). However, this is a difficult system of equations to solve, because of the time-dependent rates $\lambda(\tau)$ and $\mu(\tau)$, the time-dependent booking limit $CAP(\tau)$, and the complex nature of the upper bound on the state space $\max(B(\tau), CAP(\tau))$. It may be possible to make some approximations and assumptions to obtain a less complex system of equations, as was done in Chapter 4. We leave this issue to future research.

8.2.2 Extensions of the Statistical Models

At the end of Chapter 7, we briefly reviewed some extensions to the censored Poisson model and the full information combined model. Additionally, we proposed a new combined censored Poisson/full information model. Extensions to the censored Poisson model include characterizing the request rate $\lambda(\tau)$ and/or the cancellation rate $\mu(\tau)$ as a function of explanatory variables. Although the resulting model would be more time consuming to estimate, the potential benefit may be substantially better forecasts. A second issue of further research pertaining to the censored Poisson model is how to determine the optimal number of subintervals on which to estimate the parameters. We expect that the number of subintervals will depend on the fare class and market under consideration.

Extensions to the full information combined model include using a booking curve model such as a piecewise linear approximation to estimate the bookings to come term. Also, future research should investigate the possibility of including additional explanatory variables to the full information combined model. There is a fundamental tradeoff between potential increased forecasting accuracy from additional variables and the possibility of overfitting the data and producing poor forecasts. As with the censored Poisson model, a final issue of future research is how to determine the optimal number of subintervals on which to estimate the model.

Finally, since the censored Poisson and the full information combine models produced accurate forecasts in the case studies of Chapter 7, we propose to combined the two models in order to produce a superior forecasting model. A potential combined formulation includes embedding the full information model within the censored Poisson model by characterizing the request rate as a function of explanatory variables. A second formulation embeds the censored Poisson model within the general formulation of the full information model by substituting the censored Poisson model for the bookings to come term. To implement the second method, we propose a three step method: estimate the censored Poisson model, substitute the parameters into the bookings to come term, and, then, estimate the parameters of the full information model. Overall, a combination of the two models is an intuitively appealing idea for producing more accurate forecasts.

8.2.3 Extension to Origin-Destination Forecasting

The emphasis in this dissertation is on forecasting total bookings in a specific fare class on a particular flight leg. Traditional seat inventory control deals primarily with the problem of optimally forecasting and controlling total bookings on each flight leg. The result is that, if a seat is available in Q class, it is available to any passenger no matter what the destination, itinerary, and total fare paid. Because of the development of major hub-and-spoke networks, a flight leg from Boston to Minneapolis is inevitably carrying many passengers whose final destination is beyond Minneapolis. Airlines are starting to understand that the Boston-Tokyo Q class passenger paying $1000 is potentially much more valuable than the Boston-Los Angeles Q class passenger paying $200. Hence, there is a great amount of interest in controlling seats in an origin-destination, fare class environment.

The issue of forecasting in an origin-destination, fare class environment is not an easy issue to address. A key issue to be addressed is the "small numbers" problem. When looking at a typical flight leg into a major hub, there are many final destinations with only a very few passengers. Some final destinations may only have 1 or 2 passengers in a specific fare class on any particular incoming flight. Other final destinations may have no passengers in a specific fare class on a particular incoming flight. Thus, the issue is whether or not we can predict such small numbers.

There are two potential methods which can be used to estimate origin-destination demand. First, the small numbers associated with origin-destination demand can be considered as "count" data. The Poisson distribution assumption is usually suggested as an ideal way to analyze this type of data. The censored Poisson model developed in Chapter 4 can be directly applied to this problem. Practically, the computational requirements of the censored Poisson model might be large. Thus, a Poisson regression model based on the framework of Chapter 5 could easily be developed to forecast origin-destination demand.

A second approach to forecasting of origin-destination demand is a two-stage method. The first stage is to predict total bookings in the specific fare class on the flight leg under consideration. Any of the models discussed in this thesis can be used in the first stage. The second stage is to estimate a market share model which splits the flight leg demand into origin-destination pairs. Perhaps a logit model can be used for the market share model. The forecast of origin-destination demand is simply the product of the flight leg demand forecast and the market share forecast for the particular origin-destination under consideration.

Both of these methods may be fairly time consuming, requiring the use of maximum likelihood estimation. However, the first method requires a separate estimation for each available origin-destination pair in a particular fare class. The second method only requires two estimations in order to predict demand for all available origin-destination pairs. Further empirical and theoretical research is needed to explore the origin-destination demand forecasting problem.

8.2.4 The Need for a Forecasting System

This thesis proposes several probabilistic and statistical models of the airline booking process. The final step to practical implementation of the models described in this thesis is the development of a successful, automated forecasting system. In this section, we conceptually outline the major elements of an automated forecasting system. The details are left to future research.

The goal of the automated forecasting system is to accurately forecast total bookings by flight leg (or possibly origin-destination), class, and departure date at various points before a flight departs. To accomplish this task, we propose five modules: the data gathering module, the data preparation module, the estimation module, the forecasting module, and the performance module. A flow chart of these modules is given in Figure 8.1.

The first module is the data gathering module. The main purpose of this module is to bring together the relevant data needed for estimation and forecasting from various airline data bases. The types of data required are outlined in Chapter 6. The second module is the data

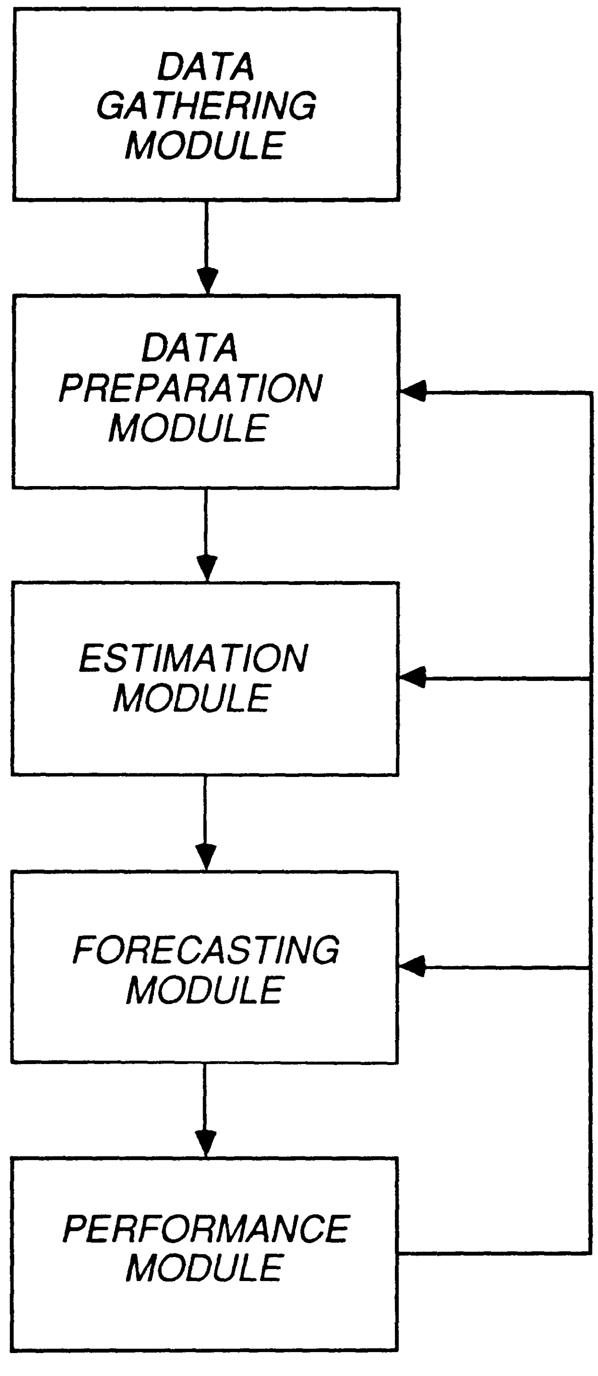

Figure 8.1 The Proposed Forecasting System

preparation module. This module prepares the data for use in the estimation and forecasting procedures. Two important elements of the data preparation module are detection and removal of "gross" outliers in the data set and elimination of the seasonal variation in the data.

The next module is the estimation module, which selects the best model and applies the appropriate estimation procedure to the data set. It is necessary to develop a good model selection procedure for this module. Model selection should not only focus on which type of model to use, but also on editing outliers and removal of variables with insignificant t-ratios and wrong signs. The fourth module is the forecasting module. This module intelligently extrapolates the estimated models to the future flight under consideration. Two key elements of this module are incorporation of procedures which attempt to reduce "bad" forecasts and possible methods for combining forecasts from different models.

The final module is the performance module. The main objective of this module is to monitor and compare the accuracy of forecasting over time as well as identify "problem" flights. This module could include running tallies of forecast performance for each flight number and identify those flights with large or growing errors. The result could be used to generate exception reports for human intervention. Ideally, the performance module would feedback directly into the earlier modules in order to enhance forecast accuracy.

In conclusion, this proposed forecasting system is simply a conceptual model. Many details remain for further research. In this thesis, we have proposed a probabilistic model and a rigorous statistical framework, addressed important issues of estimation and forecasting, tested the forecasting accuracy of two key models on actual airline data, and evaluated the worth of accurate forecasts to the airlines. The key to practical implementation of these models will be the development of a successful forecasting system.

Bibliography

Amemiya, T. (1973). "Regression Analysis When the Dependent Variable Is Truncated Normal". Econometrica, Volume 41, p. 997.

Bailey, Norman T.J. (1964). The Elements of Stochastic Processes with Applications to the Natural Sciences. New York: Wiley.

Bartlett, M.S. (1966). An Introduction to Stochastic Processes. Cambridge, England: Cambridge Press.

Beaumont, G.P. (1983). Introductory Applied Probability. New York: Wiley.

Belobaba, Peter (1985). "TWA Reservations Analysis: Demand Distribution Patterns". MIT Flight Transportation Lab Report. Cambridge, MA.: Massachusetts Institute of Technology.

Belobaba, Peter (1987). "Air Travel Demand and Airline Seat Inventory Management". MIT Thesis. Cambridge, MA.: MIT Flight Transportation Lab.

Belsey, D.A. et. al. (1980). Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. New York: Wiley.

Ben-Akiva, Moshe and D. Bolduc (1987). "Approaches to Model Transferability and Updating: The Combined Transfer Estimator". Transportation Research Record, Volume 1139, p. 1.

Ben-Akiva, Moshe, and S. Lerman (1985). Discrete Choice Analysis: Theory and Application to Travel Demand. Cambridge, MA.: MIT Press.

Ben-Akiva, Moshe. et. al. (1987). "Improving Airline Passenger Forecasts Using Reservation Data". Presentation at Fall ORSA/TIMS Conference, St. Louis, MO.

Berndt, E.R. et. al. (1974). "Estimation and Inference in Non-linear Structural Models". Annals of Economic and Social Measurement, Volume 3, Number 4, p. 653.

Bharucha-Reid, A.T. (1960). Elements of the Theory of Markov Processes and Their Applications. New York: McGraw-Hill.

Boeing Commercial Airplane Company (1982). "Boeing Promotional Fare Management System: Analysis and Research Findings". Seattle, WA.: Boeing Sales Technology Department.

Brummer, Mark et. al. (1988). "Determination of Potential Load Capacity Based on Observed Load Factor Data: A Study for Northwest Airlines". Northfield, MN.: St. Olaf College Undergraduate Practicum Group.

Chatterjee, S. and A.S. Hadi (1988). Sensitivity Analysis in Linear Regression. New York: Wiley.

Chiang, C.L. (1968). Introduction to Stochastic Processes in Biostatistics. New York: Wiley.

Cox, D.R. and H.D. Miller (1965). The Theory of Stochastic Processes. New York: Wiley.

Davis, J.E. (1987). "The Competitive Impact of Carrier Owned Computer Reservations Systems". Unpublished MIT Term Paper. Cambridge, MA.: MIT Flight Transportation Laboratory.

Ford, Lester (1955). Differential Equations, New York: McGraw-Hill.

GAUSS System Version 2.0 (1988). Kent, WA.: Aptech Systems, Inc.

Goel, N.S. and N. Richter-Dyn (1974). Stochastic Models in Biology. New York: Academic Press.

Greene, W.H. (1990). Econometric Analysis. New York: MacMillan.

Judge, George G. et. al. (1985). The Theory and Practice of Econometrics. New York: Wiley.

Kanafani, A.K. (1983). Transportation Demand Analysis. New York: McGraw-Hill.

Kells, Lyman (1975). Elementary Differential Equations. New York: McGraw-Hill.

Kennedy, Peter (1987). A Guide to Econometrics. Cambridge, MA.: MIT Press.

Larsen, R.J. and M.L. Marx (1986). An Introduction to Mathematical Statistics and its Applications. Englewood Cliffs, N.J.: Prentice-Hall.

Larson, R.C. and A.R. Odoni (1981). Urban Operations Research. Englewood Cliffs, N.J.: Prentice-Hall.

Lee, Anthony (1988). "Airline Reservations Forecasting". Proceedings of the 28th Annual AGIFORS Symposium.

Littlewood, Kenneth (1972). "Forecasting and Control of Passenger Bookings". Proceedings of the 12th Annual AGIFORS Symposium.

Maddala, G.S. (1983). Limited-Dependent and Qualitative Variables in Econometrics. Cambridge, England: Cambridge University Press.

McCullagh, P. and J.A. Nelder (1989). Generalized Linear Models. New York: Chapman and Hall.

Medhi, J. (1982). Stochastic Processes. New York: Wiley.

Montgomery, D.C. and E.A. Peck (1982). Introduction to Linear Regression Analysis. New York: Wiley.

Parzen, Emanuel (1962). Stochastic Processes. San Francisco: Holden Day.

Pindyck, R.S. and D.L. Rubinfeld (1981). Econometric Models and Economic Forecasts. New York: McGraw-Hill.

Power, C. (1990). "It's Already a Year to Forget: Fares Won't Rise Much -- and Costs Won't Drop". Business Week, January 8, 1990, p. 88.

Ricciardi, L.M. (1986). "Stochastic Population Theory: Birth and Death Processes". In T.G. Hallam and S.A. Levin (eds.), Biomathematics, Volume 17, Mathematical Ecology. New York: Springer-Verlag.

Riordan, John (1962). Stochastic Service Systems, New York: Wiley.

Rothstein, M. (1971). "An Airline Overbooking Model". Transportation Science, Volume 5, p. 180.

Rothstein, M. (1985). "OR and the Airline Overbooking Problem". Operations Research, Volume 33, p.237.

Sa, Joao (1987). "Reservations Forecasting In Airline Yield Management". MIT Thesis. Cambridge, MA.: Flight Transportation Lab.

SAS Statistical Analysis System Version 5 (1985). Cary, NC.: SAS Institute, Inc.

Scandinavian Airlines System (1978). "SHOPS: A Short-term Passenger Forecasting System".

Schneider, H. (1986). Truncated and Censored Samples from Normal Populations. New York: Marcel Dekker.

Simpson, R.W. (1982). "Notes for Air Transportation". MIT Flight Transportation Laboratory Lecture Notes. Cambridge, MA.: Massachusetts Institute of Technology.

Smith, Barry. and C. Penn (1988). "Analysis of Alternate Origin-Destination Control Strategies". Proceedings of the 28th Annual AGIFORS Symposium.

Taib, I. (1990). "Loophole Allows Bias in Displays on Computer Reservations Systems". Aviation Week, Volume 132, February 12, 1990, p. 137.

Taneja, N.K. (1978). Airline Traffic Forecasting. Lexington, MA.: Lexington Books.

Taylor, C.J. (1962). "The Determination of Passenger Booking Levels". Proceedings of the 2nd Annual AGIFORS Symposium.

Tobin, J. (1958). "Estimation of Relationships for Limited Dependent Variables". Econometrica, Volume 26, p. 24.

Williamson, E. (1988). "Comparison of Optimization Techniques for Origin-Destination Seat Inventory Control". MIT Thesis. Cambridge, MA.: Flight Transportation Laboratory.

Appendix A The Value of Forecasting in Airline Seat Inventory Control

A.1 Introduction

The goal of this appendix is to describe a simulation methodology to measure the value of more accurate forecasts of the true, unconstrained demand for total bookings. That is, the objective is to determine the impact of more accurate forecasts on expected revenues. The first section outlines the methodology used to evaluate the impact of forecasting. Next, we introduce the optimization model used to produce the nested booking limits on each fare class. The third section describes the simulation of the booking process in detail. In the fourth section, we apply the methodology to airline data and demonstrate the results. Finally, we draw conclusions from the results. The key result is that more accurate forecasts bring about an increase in expected revenues.

A.2 Overview of Methodology

This section develops a simulation model of the booking process to measure the impact of more accurate forecasts on expected revenues. This model assumes that there is a forecast of total bookings in each fare class obtained from, for example, one of the statistical models in

Chapter 5. Additionally, it is assumed that the actual distribution of total bookings (usually different from the forecast of total bookings) is known. We should note that this simulation model is *static*. That is, demand for each class arrives at a single point in time and the corresponding booking limits are calculated only once during the booking process. The goal of this simulation model is to systematically vary the forecast of total bookings relative to the actual distribution of total bookings in each fare class and determine the effect on expected revenues.

There are two key steps in the simulation model. First, we input the forecasts of total bookings in each fare class into the optimization model in order to generate the nested booking limits. Then, applying these booking limits, we perform a Monte Carlo simulation of the booking process where the total bookings in each fare class are drawn from the actual distribution of demand. As we systematically vary the forecast demand relative to the actual demand, we measure the change in expected revenues. This measure of change represents the "value" of accurate forecasting.

A.3 The Optimization Model for Nested Booking Limits

The Expected Marginal Seat Revenue (EMSR) model is a seat inventory control model originally developed by Belobaba (1987). The EMSR model is a probabilistic revenue maximization heuristic used to calculate the nested booking limits on each fare class. The aim of the EMSR model is to limit bookings in lower fare classes while protecting seats for expected demand in higher fare classes. The crucial element of the EMSR model is that seats are protected for a given fare class by equating *the expected marginal revenue of protecting an* *additional seat in this fare class with the expected marginal revenue of not protecting the seat and selling it in a lower fare class.* (Williamson, 1988).

We briefly review the mathematics of the EMSR model as given by Belobaba (1987). Let $p_i(q_i)$ be the probability that q_i or more requests are received for class i on a particular flight. We omit the flight f and departure date d subscripts for clarity. The expected number of bookings received in class i, $B_i(q_i)$, is a function of the seats made available in class i, q_i. That is, the bookings will be limited by q_i.

If we denote F_i as the average fare for class i, then the expected total revenue from having q_i seats available to class i is

$$R_i(q_i) = F_i * B_i(q_i)$$

In seat inventory control, the focus is on the expected *marginal* revenue gained from selling one more seat in class i ($EMSR_i$). Thus, according to Belobaba (1987), the expected marginal revenue for the qth seat made available to class i is the fare in class i times the probability of selling q_i or more seats in class i:

$$EMSR_i(q_i) = F_i * p_i(q_i)$$

To find the number of seats which should be protected for class i and not be made available to class j, SP_{ij}, we set the expected marginal revenue of the SP_{ij}th seat in class i equal to the fare in class j. Mathematically, we have

$$EMSR_i(SP_{ij}) = F_i * p_i(SP_{ij}) = F_j \tag{A.1}$$

When equation (A.1) holds, the airline is indifferent between the certain revenue F_j in class j and the expected revenue $EMSR_i(SP_{ij})$ for an additional seat in class i. So, for each fare class i, we find optimal protection levels SP_{ij} for all $j > i$.

Once the protection levels are determined, we must specify booking limits on each fare class. Booking limits on a particular fare class are the maximum number of seats made available to the given class and lower classes. In the EMSR framework, the booking limit for class i, CAP_i, is the capacity of the aircraft minus the seats protected for higher fare classes. Mathematically,

$$CAP_i = TCAP - \sum_{j<i} SP_{ji}$$

If we assume a probability distribution of demand, the booking limits and optimal seat protections for each fare class can be calculated. In the simulation model, we assume that the forecasts of total bookings for each fare class i are Normally distributed with mean $\hat{\mu}_i$ and standard deviation $\hat{\sigma}_i$. To determine the formula for the number of seats protected for each class under the Normal distribution assumption, we rewrite equation (A.1) as

$$1 - \Phi((SP_{ij} - \hat{\mu}_i) / \hat{\sigma}_i) = F_j / F_i \quad \text{for all } j > i \tag{A.2}$$

where Φ is the standard Normal cumulative distribution function. Now, we need to solve equation (A.2) for the number of seats protected for class i from class j, SP_{ij}. Since the fare ratio in equation (A.2) is F_j / F_i, we must find the value which has a probability of F_j / F_i of being exceeded. Tables of the standard Normal cumulative distribution function are found in most elementary statistics texts. Denote the appropriate argument of Φ by Z_{ij}. Then, we solve the following formula for SP_{ij}

$$Z_{ij} = (SP_{ij} - \hat{\mu}_i) / \hat{\sigma}_i \tag{A.3}$$

or, rearranging terms,

$$SP_{ij} = \hat{\mu}_i + (Z_{ij} * \hat{\sigma}_i) \tag{A.4}$$

Finally, we investigate the sensitivity of the number of seats protected to changes in the forecast mean and standard deviation of total bookings. This is important since the simulation model examines the impact of changes in the forecast mean and standard deviation on the booking limits and the resulting expected revenues. Equation (A.4) shows that a unit change in the forecast mean ($\hat{\mu}_i$) causes a one unit change in the number of seats protected for class i from class j (SP_{ij}) -- all else held constant.

On the other hand, the impact of changes in the forecast standard deviation ($\hat{\sigma}_i$) depends on the value of Z_{ij} and, thus, the ratio of the fares in class i and j (F_j / F_i). From the properties of the standard Normal distribution, we can deduce the following relationships:

If $0.0 < F_j / F_i \leq 0.16$, then $Z_{ij} \geq 1.0$

If $0.16 < F_j / F_i < 0.5$, then $0.0 < Z_{ij} < 1.0$

If $F_j / F_i = 0.5$, then $Z_{ij} = 0.0$

If $0.5 < F_j / F_i < 0.84$, then $-1.0 < Z_{ij} < 0.0$

If $0.84 \leq F_j / F_i < 1.0$, then $Z_{ij} \leq -1.0$

Therefore, for fare ratios between 0.16 and 0.84, $|Z| < 1.0$. In this case, a unit change in the forecast standard deviation brings about less than a unit change (in absolute value) in the number of seats protected. However, if the fares in classes i and j are nearly the same ($F_j / F_i \geq 0.84$) or are quite different ($F_j / F_i \leq 0.16$), a unit change in the forecast standard deviation causes more than a unit change in the number of seats protected.

Overall, changes in the forecast mean have a direct impact on the seat protection levels. As the forecast mean changes, the seat protection levels change by the same amount. Changes in the forecast standard deviation have a less direct impact on seat protection levels. For fare ratios between 0.16 and 0.84, a change of one unit in the forecast standard deviation produces a fractional change in the number of seats protected. Thus, for most fare levels, changes in the forecast mean have a much larger impact on seat protection levels than changes in the forecast standard deviation.

A.4 The Simulation Model

We now develop a simulation model for the booking process. The goal is to assess the impact of changes in the forecasts of total bookings relative to the actual distribution of total bookings on average revenues. The simulation is run using the SAS statistical analysis package (1985). The data for the simulation is obtained by calculating the sample mean and standard deviation of total bookings on actual flights of a major U.S. airline. Tables A.1, A.2, and A.3 summarize the data. We consider four fare classes Y, B, M, and Q with average revenues of 100, 70, 50, and 30, respectively. The number of bookings in each fare class i is assumed to be Normally distributed with mean bookings (μ_i) and an associated standard deviation (σ_i).

Since we want to measure the effect of changes in the forecast mean ($\hat{\mu}_i$) and standard deviation ($\hat{\sigma}_i$) relative to the actual mean (μ_i) and standard deviation (σ_i), we have three major cases to examine:

1. Vary the forecast standard deviation only.
2. Vary the forecast mean only.
3. Vary the forecast mean and standard deviation.

In each of the three cases, we vary the mean and/or standard deviation of each fare class by the same factor. The formula for varying the mean and/or standard deviation is stated as follows:

$$\text{Forecast} = (\text{Factor}) * \text{Actual}$$

where we vary the factor from 0.25 to 3.0. Specifically, for varying the forecast mean, the formula is

$$\hat{\mu}_i = (\text{Factor}) * \mu_i$$

For varying the forecast standard deviation, the formula is

$$\hat{\sigma}_i = (\text{Factor}) * \sigma_i$$

If the factor is greater than 1, then we overforecast the actual distribution of bookings. If the factor is less than 1, then we underforecast the actual distribution of bookings. If the factor is exactly equal to 1, then the forecast distribution is the same as the actual distribution of demand.

The simulation proceeds as follows:

Step 1: Given the forecast mean bookings ($\hat{\sigma}_i$) and forecast standard deviation ($\hat{\mu}_i$) of bookings for each fare class i, calculate the number of seats protected for each fare class using the EMSR model as presented in the previous section.

Step 2: For each fare class i, the simulation draws a random observation from a Normal distribution with actual mean μ_i and standard deviation σ_i.of bookings. This random observation is taken to be the total demand for bookings in class i. If the random observation is negative, we set it equal to zero. That is, the data is *truncated* at zero.

Step 3: We assume that demand is realized at a single point in time and that Q class passengers book first, M class passengers book next, B class passengers book third, and Y class passengers book last. Starting with Q class, demand is accepted for each class if the total bookings on the flight remain less than or equal to the corresponding booking limit. Otherwise, if demand exceeds the booking limit, the excess demand is spilled and presumed to be lost. Thus, the data is *censored* at the booking limit.

Step 4: Finally, the total revenue is determined by multiplying the accepted demand in each fare class by the average fare F_i for each fare class i.

Step 5: For each forecast of bookings, we simulate 2000 flights by repeating steps 2 through 4 for 2000 iterations. After the final iteration, average revenue (total revenue/2000) is calculated for the simulation run.

Step 6: Vary the forecast distribution of bookings for each class by a fixed factor and go back to step 1. In particular, we do the following:

- For Case 1 (vary standard deviation only), we run a simulation for $\hat{\sigma}_i$ = $(0.25^*\sigma_i)$, $(0.5^*\sigma_i)$, $(0.75^*\sigma_i)$, $(0.9^*\sigma_i)$, $(1.0^*\ \sigma_i)$, $(1.1^*\sigma_i)$, $(1.25^*\sigma_i)$, $(1.5^*\sigma_i)$, $(2.0^*\sigma_i)$, and $(3.0^*\sigma_i)$.
- For Case 2 (vary mean only), we run a simulation for $\hat{\mu}_i$ = $(0.25^*\mu_i)$, $(0.5^*\mu_i)$, $(0.75^*\mu_i)$, $(0.9^*\mu_i)$, $(0.95^*\mu_i)$, $(1.0^*\mu_i)$, $(1.05^*\mu_i)$, $(1.1^*\mu_i)$, $(1.25^*\mu_i)$, $(1.5^*\mu_i)$, $(2.0^*\mu_i)$, and $(3.0^*\mu_i)$.
- For Case 3 (vary mean and standard deviation together), we run a simulation for $(\hat{\mu}_i, \hat{\sigma}_i)$ = $(0.25^*\mu_i, 0.25^*\sigma_i)$, $(0.5^*\mu_i, 0.5^*\sigma_i)$, $(0.75^*\mu_i, 0.75^*\sigma_i)$, $(0.9^*\mu_i, 0.9^*\sigma_i)$, $(0.95^*\mu_i, 0.95^*\sigma_i)$, $(1.0^*\mu_i, 1.0^*\sigma_i)$, $(1.05^*\mu_i, 1.05^*\sigma_i)$, $(1.1^*\mu_i, 1.1^*\sigma_i)$, $(1.25^*\mu_i, 1.25^*\sigma_i)$, $(1.5^*\mu_i, 1.5^*\sigma_i)$, $(2.0^*\mu_i, 2.0^*\sigma_i)$, and $(3.0^*\mu_i, 3.0^*\sigma_i)$.

In order to fully study the effect of changes in the forecasts of total bookings in each fare class, we ran simulations for four separate demand scenarios. For each aircraft type, a simulation is performed for four separate demand scenarios. The low demand scenario has total mean demand of roughly 30% of aircraft capacity. The medium demand scenario has total mean demand of approximately 60% of capacity. The total mean demand for the high and very high scenarios is roughly 90% and 120% of capacity, respectively. The aircraft capacity was taken to be 200 seats.

A.5 Results of the Simulation Model

Our goal is to compare average expected revenues as the forecast mean and/or standard deviation varies, as aircraft size varies, and as level of demand varies. The inputs and results of the simulation model are summarized in Tables A.1 to A.3 and Figures A.1 to A.3 at the end of the Appendix. We now analyze the results for each scenario.

Case 1: Vary Forecast Standard Deviation Only

Low Demand Scenario

Examining the low demand scenario inputs and EMSR results in Table A.1, the booking limits change modestly as the forecast standard deviation varies. The greatest change is a increase of 7 in the M class booking limit, when the forecast standard deviation exceeds the actual standard deviation by 300%. Figure A.1 displays the plot of average revenue as the forecast standard deviation changes. We note that the average revenue for the low demand scenario remains *constant* as the forecast standard deviation becomes less accurate.

Medium Demand Scenario

Table A.1 shows the inputs and EMSR results for the medium demand scenario. As the forecast standard deviation changes, there are moderate changes in the booking limits. For example, when the forecast standard deviation is less than the actual standard deviation, the B and M class booking limits decrease by 4 seats and the Q class booking limit increases by 3 seats. On the other hand, when the forecast standard deviation is larger than the actual standard deviation, the B class booking limit increases by 12 spaces, the M class booking limit increases by 9 spaces, and the Q class booking limit falls by 11 spaces. In Figure A.1, average expected revenue is plotted against the forecast factor. The graph reveals *no change* in average revenue as the forecast standard deviation becomes less accurate.

High Demand Scenario

Table A.1 reveals the inputs and EMSR results for the high demand scenario. As the forecast standard deviation varies, there are moderate changes in the booking limits. For example, when the forecast standard deviation is 300% of the actual standard deviation, the B class booking limit increases by 11, the M class booking limit rises by 15, and the Q class booking limit falls by 8 spaces. Figure A.1 shows the plot of average expected revenue as the forecast standard deviation varies. As the forecast standard deviation becomes less accurate (the forecast factor changes from 1.0), the average expected revenue *decreases very slightly*. For example, when the forecast standard deviation is only 25% of the actual standard deviation, the revenue decreases by 0.11%. Conversely, when the forecast standard deviation is 300% greater than the actual standard deviation, the revenue decreases by 0.33%.

Very High Demand Scenario

The inputs and EMSR results for the very high demand scenario are shown in Table A.1. As the forecast standard deviation becomes less accurate, there are some moderate to large changes in booking limits. The largest change is in the M class booking limit which increases by 18, as the forecast standard deviation increases to 300% of the actual standard deviation. In figure A.1, the average expected revenue is plotted against the forecast factor. The general trend is a *very slight decrease* in average revenues as the forecast standard deviation changes substantially from the actual standard deviation. For example, a forecast standard deviation which is 25% of the actual standard deviation brings about a 0.21% decrease in average revenues. A 300% increase in forecast standard deviation relative to the actual causes a 0.63% decline in expected average revenues.

Conclusions for Case 1

In conclusion, varying the forecast standard deviation had very little impact on expected average revenues. In the low and medium demand scenarios, changes in the forecast standard deviation brought about no change in the expected revenues. The high and very high demand scenarios revealed extremely small decreases in the expected revenues, as the forecast standard deviation changed. Thus, the simulation model appears to show that an accurate forecast of the standard deviation of bookings is of very limited value to the seat inventory control process.

Case 2: Vary Forecast Mean Only

Low Demand Scenario

Examining the EMSR inputs and results in Table A.2, the nested booking limits for B, M, and Q classes change considerably as the forecast mean varies. For example, when the forecast mean exceeds the actual mean, the B class booking limit changes from 195 to 180, the M class booking limit falls from 183 to 144, and the Q class booking limit declines from 162 to 92. When the forecast mean falls below the actual mean, we observe more moderate changes in the booking limits. Figure A.2 shows a plot of average expected revenues as the forecast mean varies. Note that average revenue *remains essentially constant* as the forecast mean is modified. This result is expected, since the airline accepts almost all of the requested demand in the low demand scenario.

Medium Demand Scenario

Table A.2 shows the inputs and EMSR results for this scenario. The booking limits of B, M, and Q classes change significantly as the forecast mean varies. As the forecast mean exceeds the actual mean, the booking limit for B class declines from 188 to 154, for M class decreases from 169 to 98, and for Q class falls from 134 to 14. When the forecast mean is less than the actual mean, the booking limits increase moderately for B, M, and Q classes. Figure A.2 displays a plot of average expected revenues as the forecast mean changes. The graph shows *no change* in expected revenues as the forecast mean falls below the actual mean. However, when the forecast mean is greater than the actual mean, the expected revenues decline considerably. To illustrate, when the forecast mean is 300% of the actual mean, the average expected revenues decline by roughly 27%.

High Demand Scenario

The inputs and EMSR results for the high demand scenario are displayed in Table A.2. As the forecast mean varies from the actual mean, there are significant changes in the booking limits for B, M, and Q classes. The largest change in booking limits is in M class, which decreases from 147 to 28, as the forecast mean exceeds the actual mean. Figure A.2 plots average expected revenues as the forecast mean changes. For this high demand scenario, the average revenue declines significantly as the forecast mean becomes less accurate. When the forecast mean is 25% of the actual mean, average expected revenue decreases by 4.9%. On the other hand, if the forecast mean is increased by 300% over the actual mean, then average expected revenue declines by 30.2%.

Very High Demand Scenario

Table A.2 shows the inputs and EMSR results for the very high demand scenario. As the forecast mean varies, we observe large changes in the booking limits. When the forecast mean is 25% of the actual mean, the Q class booking limit increases to 166, the M class booking limit rises to 190, and the B class booking limit becomes 198. If the forecast mean is greater than the actual mean, then the B, M, and Q class booking limits decrease substantially. Figure A.2 displays the average expected revenues versus the forecast factor. The average revenue decreases considerably as the forecast mean becomes less accurate. A forecast mean which equals 25% of the actual mean causes a 14% decrease in average revenue. Conversely, when the forecast mean is 300% greater than the actual mean, there is a 39% decrease in expected average revenues.

Conclusions for Case 2

In summary, varying the forecast mean has a substantial impact on average expected revenues in the high and very high demand scenarios. No impact on expected revenues is observed in the low demand scenario. In the medium demand scenario, there is no change in average revenues when the forecast mean is smaller than the actual mean. However, when the forecast mean significantly exceeds the actual mean, the average expected revenues decline. We note that the medium, high, and very high demand scenarios exhibit larger drops in revenue when the forecast mean exceeds the actual mean than when the forecast mean is less than the actual mean.

Case 3: Vary Forecast Mean and Standard Deviation

Low Demand Scenario

Examining the EMSR inputs and results in Table A.3, the nested booking limits for B, M, and Q classes change considerably as the forecast mean and standard deviation varies. For example, when the forecast mean and standard deviation exceed the actual mean and standard deviation, the B class booking limit changes from 195 to 185, the M class booking limit falls from 183 to 151, and the Q class booking limit declines from 162 to 89. When the forecast mean and standard deviation fall below the actual values, we observe more moderate changes in the booking limits. Figure A.3 shows a plot of average expected revenues as the forecast mean and standard deviation vary. Note that average revenue *remains essentially constant* as the forecast mean and standard deviation is modified. As noted before, this result occurs because the airline accepts almost all of the requested demand in the low demand scenario.

Medium Demand Scenario

Table A.3 shows the inputs and EMSR results for this scenario. The booking limits of B, M, and Q classes change significantly as the forecast mean and standard deviation vary. As the forecast mean and standard deviation are greater than the actual mean and standard deviation, the booking limit for B class declines from 188 to 166, for M class decreases from 169 to 107, and for Q class falls from 134 to 3. When the forecast mean and standard deviation are less than the actual mean and standard deviation, the booking limits increase moderately for B, M, and Q classes. Figure A.3 displays a plot of average expected revenues as the forecast mean and standard deviation change. The graph shows *no change* in expected revenues as the forecast mean and standard deviation fall below the actual mean and standard deviation. However, as the forecast mean and standard deviation exceed the actual mean and standard deviation, the expected revenues decline considerably. For instance, when the forecast mean and standard deviation is 300% of the actual values, the average expected revenues decline by 32.2%.

High Demand Scenario

The inputs and EMSR results for the high demand scenario are displayed in Table A.3. As the forecast mean and standard deviation vary from the actual mean and standard deviation, there are significant changes in the booking limits for B, M, and Q classes. The B class booking limit decreases from 177 to 132; the M class booking limit decreases from 147 to 43; and the Q class booking limit declines from 100 to 0, when the forecast mean and standard deviation are greater than the actual mean and standard deviation. Figure A.3 plots average expected revenues as the forecast mean and standard deviation changes. For this high demand scenario, the average revenue declines significantly as the forecast mean and standard deviation become less accurate. When the forecast mean and standard deviation are only 25%

of the actual values, the result is a 4.5% decrease in average revenue. On the other hand, if the forecast mean and standard deviation is increased by 300% over the actual values, then the average expected revenue declines by 26.2%.

Very High Demand Scenario

Table A.3 shows the inputs and EMSR results for the very high demand scenario. As the forecast mean and standard deviation varies, we observe large changes in the booking limits. When the forecast mean and standard deviation are less than the actual mean and standard deviation, the Q class booking limit increases from 81 to 169, the M class booking limit rises from 137 to 183, and the B class booking limit increases from 173 to 193. If the forecast mean and standard deviation exceed the actual mean and standard deviation, then the B, M, and Q class booking limits decrease substantially. Figure A.3 displays the average expected revenues versus the forecast factor. The average revenue decreases considerably as the forecast mean and standard deviation becomes less accurate. When the forecast mean and standard deviation is only 25% of the actual mean and standard deviation, the result is a 12.8% decrease in average revenue. Conversely, when the forecast mean and standard deviation are 300% greater than the actual, there is a 32.2% decrease in expected average revenues.

Conclusions for Case 3

Overall, varying the forecast mean and standard deviation has a substantial impact on average expected revenues in the high and very high demand scenarios. In the medium demand scenario, underforecasting the actual mean and standard deviation has no effect on expected revenues. However, substantial overforecasting of the actual values leads to a considerable decline in revenues. In the low demand scenario, changes in the forecast mean and standard deviation have no impact on the average expected revenues. Again, we observe that overforecasting has more impact on average revenues than underforecasting. Generally, case 3 (varying the mean and standard deviation) results are almost identical to case 2 (varying the mean only) results. In most cases, the differences in average expected revenues are less than 2%. When the forecast values exceed the actual by 300%, the differences in average expected revenues between cases 2 and 3 range from 4% to 7%.

A.7 Conclusions

In this section, we will compare the results from Cases 1, 2, and 3. Second, we examine the effect of underforecasting versus overforecasting. Then, two potential qualifications to the results of the simulation model are discussed. Finally, we make some general observations and conclusions.

When the results are compared across Cases 1, 2, and 3, it becomes clear that the forecast mean is the driving force of the forecasting and seat optimization models. In Case 1 (varying the forecast standard deviation only), changes in the forecast standard deviation have very little effect on average expected revenues. Case 2 (varying the forecast mean only) demonstrates that changes in the forecast mean can have a significant impact on expected revenues, particularly in the high and very high demand scenarios. Case 3 (varying the forecast mean and standard deviation only) gives very similar results to case 2. Thus, it seems that our forecasting efforts should focus on obtaining a good point estimate of the mean forecast.

Second, in each of the cases, we observe that overforecasting has a far greater impact on expected revenues than does underforecasting. The disparity between underforecasting and overforecasting is expected. Underforecasting simply means fewer seats protected for the higher fare classes and, hence, some seats are sold at lower fares when they could have been sold at higher fares. This leads to a modest loss of average expected revenue. Overforecasting results in large numbers of seats protected for the higher fare classes and, thus, few seats available at lower fares. Then, much of the actual low fare demand is lost and the forecasted high fare demand does not show up. This can lead to a large loss in expected revenue. Hence, we should avoid gross overforecasting, particularly in medium, high, and very high demand scenarios.

There are four qualifications that we must append to the results of this simulation of the value of forecasting. First, we assume that all demand occurs at a single point in time and class by class from lowest fare to highest fare class. Obviously, this is not true in actual airline operations. In reality, fare class demand is interspersed and occurs over a period of weeks and months before a flight departs. Therefore, the airline has time to update demand forecasts and booking limits during the time before departure of a particular flight. This would decrease the number of "gross" errors made by the forecasting and optimization models. Thus, the decreases in expected revenues associated with forecasting errors are overstated in this simulation model. However, the results give us an upper bound on the changes in expected revenues. Dynamic updating of forecasts and booking limits can be incorporated in future simulation models.

The second qualification is that we assume that the forecast means and standard deviations of all fare classes are varied in the same manner. Clearly, this is not always the case. It may well happen that the forecast of Y class demand exceeds the actual demand and the forecast of Q class demand is less than the actual demand, for instance. If differing variations in forecast demand are required, this is a potential future modification to the simulation model. The third qualification is the fare levels utilized in the simulation model. Although the relative fares used in the simulation model are typical for a major airline, we may desire to vary the fares to take different types of markets into account. Thus, we may be able to pinpoint types of markets where increased forecast accuracy is particularly important.

The final qualification is that generally we do not expect huge forecast errors, such as 25% of the actual value or 300% of the actual value. However, our empirical evidence shows that a forecast between 50% and 150% of the actual value is common. Thus, in our conclusions, we focus on this range of forecast errors.

From the standpoint of average expected revenues, increased accuracy of the forecast mean *brings about a significant increase in average revenues* in the high and very high demand scenarios. On the other hand, in the low and medium demand scenarios, increases in forecast accuracy do not generally bring about a significant increase in average revenues. We conclude, then, from this simulation that improved forecasting is crucial on high and very high demand flights. On these flights, "every seat counts". On high and very high demand flights, there are increases of 0.5% to 3% in average revenues for every 10% increase in forecast accuracy.

We can demonstrate the value of accurate forecasts by examining data from a major U.S. air carrier. A sample of the carrier's flight leg data shows that an average of 16% of the flight legs have a load factor of 90% or more. These flight legs roughly correspond to our definition of high and very high demand flights. An average of 2373 flight legs were operated each day during the sample period. Thus, the average number of high and very high demand flight legs is 2373 * 16% = 380. The average number of passengers per flight leg on high and very high demand flights is 135 seats. Hence, the average number of passengers per day on high and very high demand flights is 380 * 135 = 51300. Finally, the average fare paid per passenger per flight leg is $105. Therefore, the average revenue per day on high and very high demand flights is 51300 * 105 = $5,386,500.

On an annual basis, the average revenue on high and very high demand flight legs is $5,386,500 * 365 = $1.966 billion. The simulation results indicate that each 10% of forecast accuracy brings about a 0.5% to 3.0% increase in average revenues on high and very high demand flights. Thus, for a major U.S. airline, each 10% of forecast accuracy can mean from $9.8 million to $58.8 million per year in increased revenue - - a significant amount of revenue.

TABLES A.1 - A.3

FIGURES A.1 - A.3

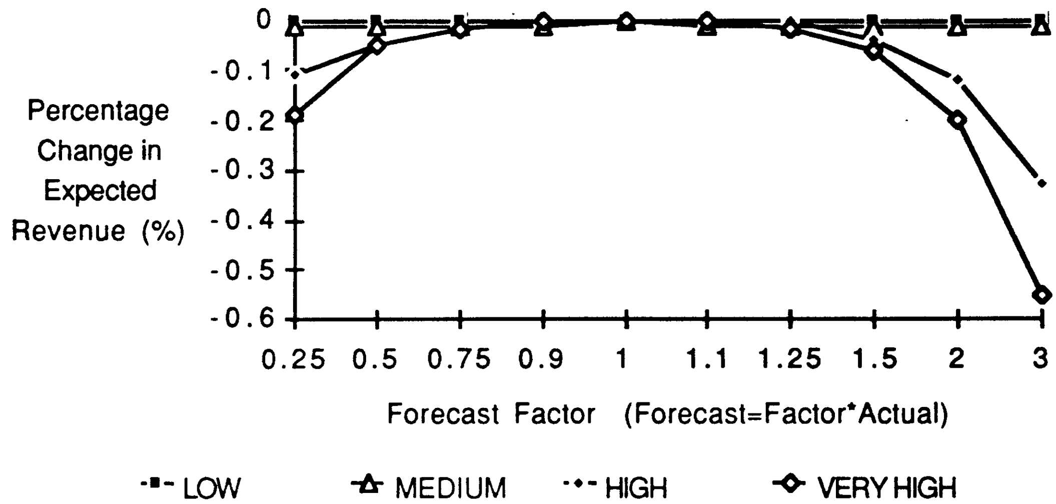

Figure A.1: Percentage Change in Average Expected Revenue Varying the Forecast Standard Deviation Only -- Capacity of 200

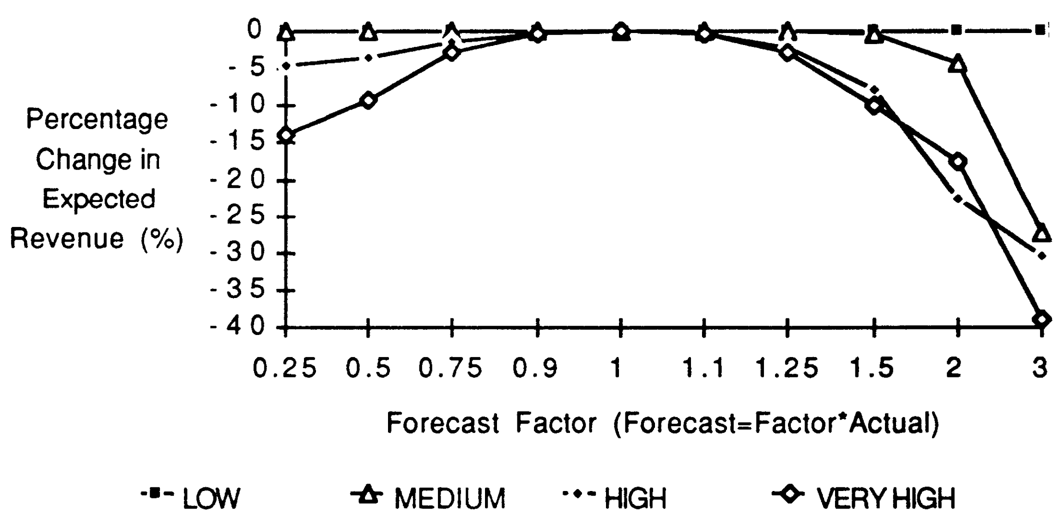

Figure A.2: Percentage Change in Average Expected Revenue Varying the Forecast Mean Only -- Capacity of 200

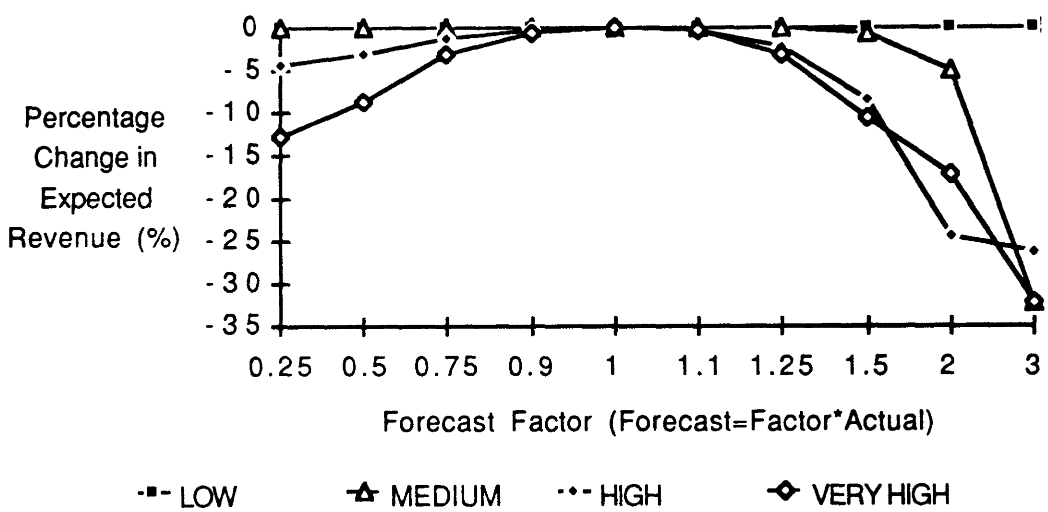

Figure A.3: Percentage Change in Average Expected Revenue Varying the Forecast Mean and Standard Deviation -- Capacity of 200

Table A.1 -- Inputs and EMSR Results for 200 Seat Aircraft
Case 1: Vary Standard Deviation Only

	Y	B	M	Q
Average Revenue	100	70	50	30
LOW DEMAND SCENARIO				
Actual:				
Actual Mean Bookings	7.4	12.2	15.6	28.1
Actual Standard Deviation of Bookings	4.9	6.8	9.8	10.1
EMSR Booking Limits	200	195	183	162
Forecast Standard Deviation		EMSR Booking Limits		
25% of Actual Standard Deviation	200	193	180	163
50% of Actual Standard Deviation	200	193	181	163
75% of Actual Standard Deviation	200	194	182	162
90% of Actual Standard Deviation	200	194	183	162
ACTUAL	200	195	183	162
110% of Actual Standard Deviation	200	195	184	162
125% of Actual Standard Deviation	200	195	184	162
150% of Actual Standard Deviation	200	196	185	161
200% of Actual Standard Deviation	200	197	187	161
300% of Actual Standard Deviation	200	200	191	159
MEDIUM DEMAND SCENARIO				
Actual:				
Actual Mean Bookings	17.0	18.5	24.6	72.4
Actual Standard Deviation of Bookings	10.8	8.0	8.6	25.0
EMSR Booking Limits	200	188	169	134
Forecast Standard Deviation		EMSR Booking Limits		
25% of Actual Standard Deviation	200	184	165	137
50% of Actual Standard Deviation	200	185	166	136
75% of Actual Standard Deviation	200	187	167	135
90% of Actual Standard Deviation	200	188	168	134
ACTUAL	200	188	169	134
110% of Actual Standard Deviation	200	189	169	132
125% of Actual Standard Deviation	200	190	170	132
150% of Actual Standard Deviation	200	191	171	131
200% of Actual Standard Deviation	200	194	173	128
300% of Actual Standard Deviation	200	200	178	123

Table A.1 (continued)-- Inputs and EMSR Results for 200 Seat Aircraft
Case 1: Vary Standard Deviation Only

HIGH DEMAND SCENARIO

Actual:				
Actual Mean Bookings	28.4	30.8	34.5	100.5
Actual Standard Deviation of Bookings	11.0	13.0	14.5	38.6
EMSR Booking Limits	200	177	147	100
Forecast Standard Deviation		EMSR Booking Limits		
25% of Actual Standard Deviation	200	173	142	104
50% of Actual Standard Deviation	200	174	143	103
75% of Actual Standard Deviation	200	175	145	102
90% of Actual Standard Deviation	200	176	146	101
ACTUAL	200	177	147	100
110% of Actual Standard Deviation	200	177	148	100
125% of Actual Standard Deviation	200	178	149	100
150% of Actual Standard Deviation	200	180	151	98
200% of Actual Standard Deviation	200	183	154	96
300% of Actual Standard Deviation	200	188	162	92

VERY HIGH DEMAND SCENARIO

Actual:				
Actual Mean Bookings	33.6	36.7	43.0	125.5
Actual Standard Deviation of Bookings	13.8	15.3	20.5	48.2
EMSR Booking Limits	200	173	137	81
Forecast Standard Deviation		EMSR Booking Limits		
25% of Actual Standard Deviation	200	168	131	84
50% of Actual Standard Deviation	200	170	133	82
75% of Actual Standard Deviation	200	171	135	81
90% of Actual Standard Deviation	200	172	137	80
ACTUAL	200	173	137	81
110% of Actual Standard Deviation	200	174	138	80
125% of Actual Standard Deviation	200	175	140	79
150% of Actual Standard Deviation	200	177	142	78
200% of Actual Standard Deviation	200	180	146	75
300% of Actual Standard Deviation	200	188	155	71

Table A.2 -- Inputs and EMSR Results for 200 Seat Aircraft
Case 2: Vary Mean Only

	Y	B	M	Q
Average Revenue	100	70	50	30
LOW DEMAND SCENARIO				
Actual:				
Actual Mean Bookings	7.4	12.2	15.6	28.1
Actual Standard Deviation of Bookings	4.9	6.8	9.8	10.1
EMSR Booking Limits	200	195	183	162
Forecast Mean		EMSR Booking Limits		
25% of Actual Mean	200	200	198	188
50% of Actual Mean	200	198	193	179
75% of Actual Mean	200	197	188	170
90% of Actual Mean	200	195	185	165
ACTUAL	200	195	183	162
110% of Actual Mean	200	194	181	159
125% of Actual Mean	200	193	178	153
150% of Actual Mean	200	191	173	145
200% of Actual Mean	200	187	164	127
300% of Actual Mean	200	180	144	92
MEDIUM DEMAND SCENARIO				
Actual:				
Actual Mean Bookings	17.0	18.5	24.6	72.4
Actual Standard Deviation of Bookings	10.8	8.0	8.6	25.0
EMSR Booking Limits	200	188	169	134
Forecast Mean		EMSR Booking Limits		
25% of Actual Mean	200	200	194	179
50% of Actual Mean	200	197	186	163
75% of Actual Mean	200	192	177	148
90% of Actual Mean	200	190	171	140
ACTUAL	200	188	169	134
110% of Actual Mean	200	186	165	128
125% of Actual Mean	200	184	159	119
150% of Actual Mean	200	180	150	103
200% of Actual Mean	200	171	133	73
300% of Actual Mean	200	154	98	14

Table A.2 (continued)-- Inputs and EMSR Results for 200 Seat Aircraft
Case 2: Vary Mean Only

HIGH DEMAND SCENARIO

Actual:				
Actual Mean Bookings	28.4	30.8	34.5	100.5
Actual Standard Deviation of Bookings	11.0	13.0	14.5	38.6
EMSR Booking Limits	200	177	147	100

Forecast Mean		EMSR Booking Limits		
25% of Actual Mean	200	198	191	171
50% of Actual Mean	200	191	176	148
75% of Actual Mean	200	184	162	123
90% of Actual Mean	200	180	153	109
ACTUAL	200	177	147	100
110% of Actual Mean	200	174	141	90
125% of Actual Mean	200	170	132	77
150% of Actual Mean	200	163	118	53
200% of Actual Mean	200	148	88	7
300% of Actual Mean	200	120	28	0

VERY HIGH DEMAND SCENARIO

Actual:				
Actual Mean Bookings	33.6	36.7	43.0	125.5
Actual Standard Deviation of Bookings	13.8	15.3	20.5	48.2
EMSR Booking Limits	200	173	137	81

Forecast Mean		EMSR Booking Limits		
25% of Actual Mean	200	198	190	166
50% of Actual Mean	200	190	173	136
75% of Actual Mean	200	182	155	108
90% of Actual Mean	200	177	144	92
ACTUAL	200	173	137	81
110% of Actual Mean	200	171	134	75
125% of Actual Mean	200	165	120	52
150% of Actual Mean	200	156	102	24
200% of Actual Mean	200	140	67	0
300% of Actual Mean	200	106	0	0

Table A.3 -- Inputs and EMSR Results for 200 Seat Aircraft
Case 3: Vary Mean and Standard Deviation

	Y	B	M	Q
Average Revenue	100	70	50	30
LOW DEMAND SCENARIO				
Actual:				
Actual Mean Bookings	7.4	12.2	15.6	28.1
Actual Standard Deviation of Bookings	4.9	6.8	9.8	10.1
EMSR Booking Limits	200	195	183	162
Forecast Mean and Standard Deviation		EMSR Booking Limits		
25% of Actual Mean and Std. Dev.	200	198	195	189
50% of Actual Mean and Std. Dev.	200	197	191	181
75% of Actual Mean and Std. Dev.	200	196	187	171
90% of Actual Mean and Std. Dev.	200	195	185	166
ACTUAL	200	195	183	162
110% of Actual Mean and Std. Dev.	200	194	181	159
125% of Actual Mean and Std. Dev.	200	193	179	153
150% of Actual Mean and Std. Dev.	200	192	175	144
200% of Actual Mean and Std. Dev.	200	190	168	126
300% of Actual Mean and Std. Dev.	200	185	151	89
MEDIUM DEMAND SCENARIO				
Actual:				
Actual Mean Bookings	17.0	18.5	24.6	72.4
Actual Standard Deviation of Bookings	10.8	8.0	8.6	25.0
EMSR Booking Limits	200	188	169	134
Forecast Mean and Standard Deviation		EMSR Booking Limits		
25% of Actual Mean and Std. Dev.	200	197	191	183
50% of Actual Mean and Std. Dev.	200	194	184	166
75% of Actual Mean and Std. Dev.	200	191	176	150
90% of Actual Mean and Std. Dev.	200	189	171	140
ACTUAL	200	188	169	134
110% of Actual Mean and Std. Dev.	200	187	165	128
125% of Actual Mean and Std. Dev.	200	185	160	117
150% of Actual Mean and Std. Dev.	200	183	153	101
200% of Actual Mean and Std. Dev.	200	177	138	69
300% of Actual Mean and Std. Dev.	200	166	107	3

Table A.3 (continued)-- Inputs and EMSR Results for 200 Seat Aircraft
Case 3: Vary Mean and Standard Deviation

HIGH DEMAND SCENARIO

Actual:				
Actual Mean Bookings	28.4	30.8	34.5	100.5
Actual Standard Deviation of Bookings	11.0	13.0	14.5	38.6
EMSR Booking Limits	200	177	147	100
Forecast Mean and Standard Deviation		EMSR Booking Limits		
25% of Actual Mean and Std. Dev.	200	194	186	174
50% of Actual Mean and Std. Dev.	200	188	173	149
75% of Actual Mean and Std. Dev.	200	183	160	125
90% of Actual Mean and Std. Dev.	200	179	152	111
ACTUAL	200	177	147	100
110% of Actual Mean and Std. Dev.	200	175	142	91
125% of Actual Mean and Std. Dev.	200	171	134	76
150% of Actual Mean and Std. Dev.	200	166	121	51
200% of Actual Mean and Std. Dev.	200	154	96	2
300% of Actual Mean and Std. Dev.	200	132	43	0

VERY HIGH DEMAND SCENARIO

Actual:				
Actual Mean Bookings	33.6	36.7	43.0	125.5
Actual Standard Deviation of Bookings	13.8	15.3	20.5	48.2
EMSR Booking Limits	200	173	137	81
Forecast Mean and Standard Deviation		EMSR Booking Limits		
25% of Actual Mean and Std. Dev.	200	193	183	169
50% of Actual Mean and Std. Dev.	200	186	168	140
75% of Actual Mean and Std. Dev.	200	180	152	110
90% of Actual Mean and Std. Dev.	200	176	143	93
ACTUAL	200	173	137	81
110% of Actual Mean and Std. Dev.	200	171	132	69
125% of Actual Mean and Std. Dev.	200	167	122	50
150% of Actual Mean and Std. Dev.	200	160	106	21
200% of Actual Mean and Std. Dev.	200	147	75	0
300% of Actual Mean and Std. Dev.	200	120	14	0